

Faculté des bioingénieurs

Monitoring the Plant Water Content in Senescent Maize at Parcel Level Using Multi-Frequency SAR Data

A case study in the Taiyuan Basin, China

Auteur	Sébastien Saelens
Promoteurs	Pierre Defourny (ELIE) Jean Bouchat (ELIE)
Lecteurs	Sophie Bontemps (ELIE) Sébastien Lambot (ELIE)
Année Académique	2023 - 2024

Mémoire de fin d'études présenté en vue de l'obtention du diplôme de
Bioingénieur en Sciences et technologies de l'environnement

Acknowledgements

Premièrement, j'aimerais exprimer ma gratitude à mon promoteur, Pierre Defourny, pour m'avoir permis de réaliser ce mémoire et m'avoir donné l'opportunité de découvrir la Chine. Merci pour son encadrement de qualité et ses conseils avisés lors de nos entrevues.

Je tiens également à remercier chaleureusement mon deuxième promoteur, Jean Bouchat. Merci pour sa compagnie tout au long du séjour à Taiyuan, tant lors des excursions de terrain que dans les moments de tourisme et de détente. Je lui suis également reconnaissant pour toute son aide et son soutien durant l'année ; sans sa présence inestimable, je n'aurais probablement pas su rendre ce mémoire en juin.

I would also like to thank our Chinese partners, Jinlong Fan, Su Qiaomei and Yuejiao Liao, for their assistance in Taiyuan and during the field campaign, and their enthusiasm in helping us discover their country. 谢谢 !

Je tiens à saluer mes lecteurs, Sophie Bontemps et Sébastien Lambot, pour les remercier d'avoir pris le temps de lire ce travail avec attention.

Un grand merci à Jérôme pour son aide en codage, ainsi qu'à Louis pour son assistance en cartographie.

Je souhaite également remercier Léna, le Turbokot, le Kap Vert, et l'option 5E, pour leur bonne humeur et leur soutien tout au long de l'année. Vous m'avez fourni un cadre de travail et de détente exceptionnel, l'obtention du diplôme sera également grâce à vous.

Merci infiniment à ma maman pour sa relecture assidue et ses nombreux conseils de rédaction.

Enfin, je remercie chaleureusement ma famille, sans qui j'aurais sans doute baissé les bras après ma première année. Merci d'avoir toujours cru en moi et de m'avoir accompagné à chaque étape de cette aventure.

Table of contents

List of figures	5
List of tables	7
List of acronyms.....	8
Chapter 1: Introduction	9
Chaper 2: Literature Review	10
1. Plant Water Content in Maize.....	10
1.1 Maize and its importance	10
1.2 Impact of level of maturity	11
1.3 Assessing the level of maturity	11
2. Remote Sensing Data for Plant Water Content Crop monitoring.....	13
2.1 Optical remote sensing	13
2.2 Synthetic Aperture Radar (SAR).....	14
Wavelength and frequency	14
Polarisation and scattering mechanisms.....	15
Speckle	18
2.3 The use of SAR for vegetation monitoring and water content detection	18
X-band.....	18
C-band	19
L-band	19
Chapter 3: Objectives	21
Chapter 4: Materials and Methods	22
1. Region of Interest	22
1.1 Site description	22
1.2 Parcels characteristics	22
2. Variables of Interest.....	25
2.1 Fresh biomass.....	25
2.2 Dry biomass.....	25
2.3 Water content.....	25
3. Field Campaign	26

3.1	Field visit protocol	26
3.2	Field variables sampling.....	27
	Planting Row Direction	27
	Plant Density	27
3.3	Crop variables sampling.....	27
	Crop Height	28
	Fresh Biomass	29
	Dry Biomass.....	29
	Water Content.....	30
4.	SAR Data.....	30
4.1	X-band from TerraSAR-X.....	30
4.2	C-band from Sentinel-1	32
4.3	Data format.....	34
5.	Data Analysis and Processing	36
Chapter 5: Results and Interpretations		37
1.	Field Campaign	37
1.1	First Field Visit (16-17 August).....	37
1.2	Second Field Visit (23-24 August)	37
1.3	Third Field Visit (31 August – 1 September)	38
1.4	Fourth Field Visit (6-7 September)	38
1.5	Fifth Field Visit (20 September).....	38
2.	Temporal Evolution of Variables of Interest	39
2.1	Fresh biomass.....	39
2.2	Dry Biomass.....	40
2.3	Water Content.....	41
2.4	Cobs and Plants Water Content	43
3.	SAR Data.....	44
3.1	TerraSAR-X.....	44
3.2	Sentinel-1	44
4.	Sensitivity Analysis	45

4.1	TerraSAR-X.....	45
4.2	Sentinel-1	46
5	Linear Correlation Analysis	47
6	Non-Linear Correlation Analysis	50
6.1	Feature Importance Scores	50
6.2	Performance metrics.....	51
6.3	GridSearchCV and Hyperparameters	52
6.4	Leave-one-out Cross-Validation.....	53
6.5	Results for WC [g] and AWC [g/m ²].....	53
6.5.1	WC [g].....	53
6.5.2	AWC [g/m ²].....	55
Chapter 6: Discussion.....		58
Chapter 7: Conclusion.....		61
References		63

List of figures

Figure 1 - Our Chinese collaborators from TYUT (in the middle), Jean Bouchat (right) and myself (left) posing between two maize fields during one of our field expeditions.....	9
Figure 2 - Different growth stages of a maize plant, including the senescent stage (R6)	12
Figure 3 - Penetration of different SAR bands into vegetation cover (SAR Handbook).....	14
Figure 4 - SAR signals are transmitted and received either vertically (V) or horizontally (H). This gives the potential for four different polarization combinations (transmit listed first, receive second): VV, VH, HH, and HV (Introduction to SAR, s. d.).....	16
Figure 5 - SAR scattering mechanisms (Meyer, 2019)	17
Figure 6 - Selected parcels within the Taiyuan Basin	23
Figure 7 - Typical average parcel structure, containing multiple narrow fields managed by various owners leading to field heterogeneity.....	24
Figure 8 - Clear case of field heterogeneity, showing significant variations in plant height within the same parcel (Parcel 4)	24
Figure 9 - Schematic sampling approach	28
Figure 10 - Illustration of the considerable maize plant height, in comparison to me (1m84)	28
Figure 11 - Jean Bouchat and our Chinese collaborators cutting maize samples using shears and placing them into mesh bags during a fieldwork session (Parcel 3)	29
Figure 12 - One of our Chinese partners measuring the dry biomass weight in the TYUT lab.....	29
Figure 13 - TSX image taken on 10 October, comprising Parcel 4 (red dot).....	31
Figure 14 - Example of TSX image not including any of the 5 parcels (21 July)	32
Figure 15 - Temporal evolution of CR for the year 2023 for all maize parcels	33
Figure 16 - Temporal evolution of the fresh biomass in grams for the five parcels	39
Figure 17 - Temporal evolution of the dry biomass in grams for the five parcels	40
Figure 18 - Dry Biomass Ratio in percentage for the five parcels (16/08 – 09/20)	41
Figure 19 - Temporal evolution of the water content for the five parcels (16/08 - 20/09).....	41
Figure 20 - Water Content Ratio [%] for the five parcels (16/08 – 09/20)	42
Figure 21 - Temporal evolution of the water content per square meter for the five parcels ...	42
Figure 22 - Temporal evolution of the water content in cobs (left) and stems and leaves (right) of selected samples for the five parcels (16/08 – 20/09).....	43
Figure 23 - TerraSAR-X Cross Ratio data points spanning the field campaign duration	44
Figure 24 - Sentinel-1 Cross Ratio data points spanning the field campaign duration	45
Figure 25 - Scatter plot for TerraSAR-X data merged with Area Water Content data.....	46
Figure 26 - Scatter plot for Sentinel-1 data merged with Area Water Content data.....	46
Figure 27 - Correlation matrix between in-situ- and Sentinel-1 data.....	47
Figure 28 - Correlation matrix between in-situ, Sentinel-1 and TerraSAR-X data.....	49
Figure 29 - Actual vs predicted WC [g] using the optimised Random Forest model	54
Figure 30 - LOOCV Results for WC [g]	55

Figure 31 - Actual vs Predicted AWC [g/m ²] using the optimised Random Forest model	56
Figure 32 - LOOCV results for AWC [g/m ²].....	57
Figure 33 - Scatter plot of S1 CR (left), VV (middle) and VH (right) backscatter with in-situ measured VWC [kg/m ²]. Corn measures are indicated in blue. (Vreugdenhil et al. (2018)) ..	58

List of tables

Table 1 - Overview of the most used frequencies and wavelengths in SAR and their application (NASA SAR Handbook)	15
Table 2 - SAR polarisations (Introduction to SAR - HyP3, s.d.)	15
Table 3 - Parcel characteristics (area, width, length, and coordinates)	23
Table 4 - Acquisition dates for in-situ measurements and radar backscatter coefficients. #S1_Days and #TSX_Days indicate the number of days between the in-situ measurements and the respective satellite passage.	26
Table 5 - TerraSAR-X available datapoint, including parcel number, date and time of acquisition and various backscatter coefficients	32
Table 6 - Sentinel-1 selected datapoints, including parcel number, date and time of acquisition and various backscatter coefficients	34
Table 7 - TSX dataset with additional radar indexes (VH and VV:mean in decibels, CR and dpRVI), including parcel number, date and time of acquisition, the platform, the orbit number, and TSX count.	35
Table 8 - S1 dataset with additional radar indexes (VH and VV:mean in decibels, CR and dpRVI), including parcel number, date and time of acquisition, the platform, the orbit number, and S1 count.	35
Table 9 - Variables collected during the first field visit (16-17 August)	37
Table 10 - Variables collected during the second field visit (23-24 August)	38
Table 11 - Variables collected during the third field visit (31 August - 1 September)	38
Table 12 - Variables collected during the fourth field visit (6-7 September)	38
Table 13 - Variables collected during the fifth field visit (20 September)	39
Table 14 - Feature Importance Score for all Sentinel-1 backscatter coefficients and indexes on WC [g] and AWC [g/m ²]	51
Table 15 - Feature Importance Scores for WC [g]	53
Table 16 - Feature Importance Scores for AWC [g/m ²]	55

List of acronyms

AWC	Area Water Content [g/m ²]
CR	Cross Ratio
dpRVI	Dual-Polarisation Vegetation Index
ELIE	Earth and Life Institute – Environmental sciences
GAI	Green Area Index
LAI	Leaf Area Index
LOOCV	Leave-One-Out Cross-validation
NDVI	Normalized Difference Vegetation Index
Oob score	Out-of-bag score
PWC	Plant Water Content
S1	Sentinel-1
SAR	Synthetic Aperture Radar
TSX	TerraSAR-X
TYUT	Taiyuan University of Technology
VOD	Vegetation Optical Depth
WC	Water Content

Chapter 1: Introduction

Maize is one of the most widely grown crops globally, playing a pivotal role in agricultural economies and food security. Monitoring the water and dry matter content of maize crops as they reach their final developmental stages is important for ensuring optimal crop yields. Predicting the optimal time of harvest can also improve harvest logistics by enabling the efficient sharing of mechanical resources among multiple small landowners. Synthetic Aperture Radar (SAR) remote sensing, due to its sensitivity to vegetation water content, has the potential to offer a cost-effective solution for large-scale monitoring of maize crops and thereby to address these points.

In August and September of 2023, an intensive field campaign was carried out in the Taiyuan Basin in North-Western China. Local support was provided by Chinese partners from the Taiyuan University of Technology (TYUT). During this campaign, biophysiological parameters such as plant density, height, wet and dry biomass and water content were measured on five selected parcels. Each parcel was visited five times, approximately one week apart, resulting in 22 observations. These efforts aimed to gather comprehensive in-situ data on the final developmental stages of maize crops.

Based on this in-situ data, this study focuses on the analysis of water content dynamics using multi-frequency SAR backscatter data. By integrating SAR remote sensing data from TerraSAR-X and Sentinel-1 satellites with the field measurements, this research aims to improve agricultural management practices and optimise resource use. This study contributes to the understanding of SAR remote sensing applications in precision agriculture, offering a method which could ultimately make the effective monitoring of maize crop conditions possible.



Figure 1 - Our Chinese collaborators from TYUT (in the middle), Jean Bouchat (right) and myself (left) posing between two maize fields during one of our field expeditions

Chaper 2: Literature Review

This chapter will concentrate on the current state of scientific knowledge and understanding related to radar remote sensing and its use in plant water content monitoring, and more precisely in maize fields.

First, the importance of maize and its maturity levels in agricultural contexts will be examined. Then, the focus will be on remote sensing technologies used to monitor agricultural crops, particularly on Synthetic Aperture Radar (SAR), encompassing aspects such as radar bands, frequency, scattering mechanisms and signal polarisation. The final part of this chapter will provide an overview of recent scientific studies related to the use of Sentinel-1 and TerraSAR-X imagery for retrieving plant water content in maize crops.

1. Plant Water Content in Maize

1.1 Maize and its importance

Maize (*Zea mays L.*), also sometimes referred to as corn, is one of the most important cereal crops in the world and has the widest geographic distribution compared to any other cereal. It is estimated that a third of all farms globally grow maize and that by 2030, maize will have overtaken wheat as the most widely grown crop by area. The biggest producers of maize in the world are the United States and China who together produce about 54% of the global maize crop (Erenstein et al., 2021).

Compared to wheat and rice, maize is a more versatile crop and can be used for multiple purposes. Its primary use is as a livestock feed crop, especially in developed economies. Silage maize and maize stover as by-product of grain production provide important sources for animal feed. Fifty-six percent of the global maize production are used to feed animals (Erenstein et al., 2022). By-products of fuel ethanol production provide a large part of cattle feed in the US (Klopfenstein et al., 2013).

Maize also plays an important role as food crop, notably in sub-Saharan Africa and Latin America, where it is an important source of carbohydrates and protein, being consumed mainly as maize flour or maize meal (Ngoune Tandzi & Mutengwa, 2020).

Finally, maize is also used in the industry and energy sectors. Maize is the base for a series of different products, such as starch, sweeteners, oil, beverages, glue, industrial alcohol and fuel ethanol (Ranum et al., 2014).

1.2 Impact of level of maturity

The stage of maturity at the time of harvest has an impact on a variety of characteristics of the harvested maize such as yield and quality of the maize grains and stover. Maize senescence plays a crucial role in determining the optimal time of harvest. Delaying harvest until the senescence stage allows the plant to accumulate more dry matter, starch and other nutrients, resulting in higher yields. During senescence, the plant also relocates nutrients from leaves to kernels which has an impact on silage quality and starch content (Zhou & Yang, 2023).

Yield is directly impacted by the level of maturity of the maize plant. Grain yield typically rises as maturity levels increase into senescence. But as the plants tend to lose leaves in later stages of maturity, stover, total crop residue and total biomass matter yield tend to decrease with increasing maturity of the plant (Tolera et al., 1998a).

Maturity level also impacts moisture content at time of harvest which has a direct impact on the ability to safely store maize. Higher moisture levels lead to a higher risk of mold and fungi development (Laca et al., 2006).

The nutritive value of maize silage for cow feed is also impacted by the level of maturity of the maize plants. The nutritive value of silage depends on digestibility and starch content, and multiple studies have shown that digestibility of the stover decreases dramatically in plant senescence when the plant reaches the latest levels of maturity (Bal et al., 1997; Johnson et al., 1999; Jensen et al., 2005). The stage of maturity has also been shown to affect ensilage properties with regards to aerobic stability (Filya, 2004). Increasing the maturity of whole plant maize at harvest is also thought to contribute to a reduction in the methane emission of the cows fed with silage of this maize (Hatew et al., 2016).

1.3 Assessing the level of maturity

The corn growth cycle can be broadly divided into two main phases: vegetative (V) and reproductive (R), as shown in Figure 1. Within the reproductive stage, there are six distinct phases which can be distinguished by kernel size and shape: silking, blistering (R2), milking, (R3), doughing (R4), denting (R5), and reaching physiological maturity (R6) as also shown in Figure 2. Physiological maturity with onset of plant senescence typically happens around 60 days after silking and roughly 20 days following denting.

Monitoring corn growth stages up to V10 is crucial for implementing effective management practices. Adjusting fertilisation, irrigation and other practices based on the observed growth stages ensures crops develop optimally during these stages. Moisture and heat stress during these stages will significantly impact final yield, so timely information about the growth stage and plant water content are vital and can improve farming techniques.

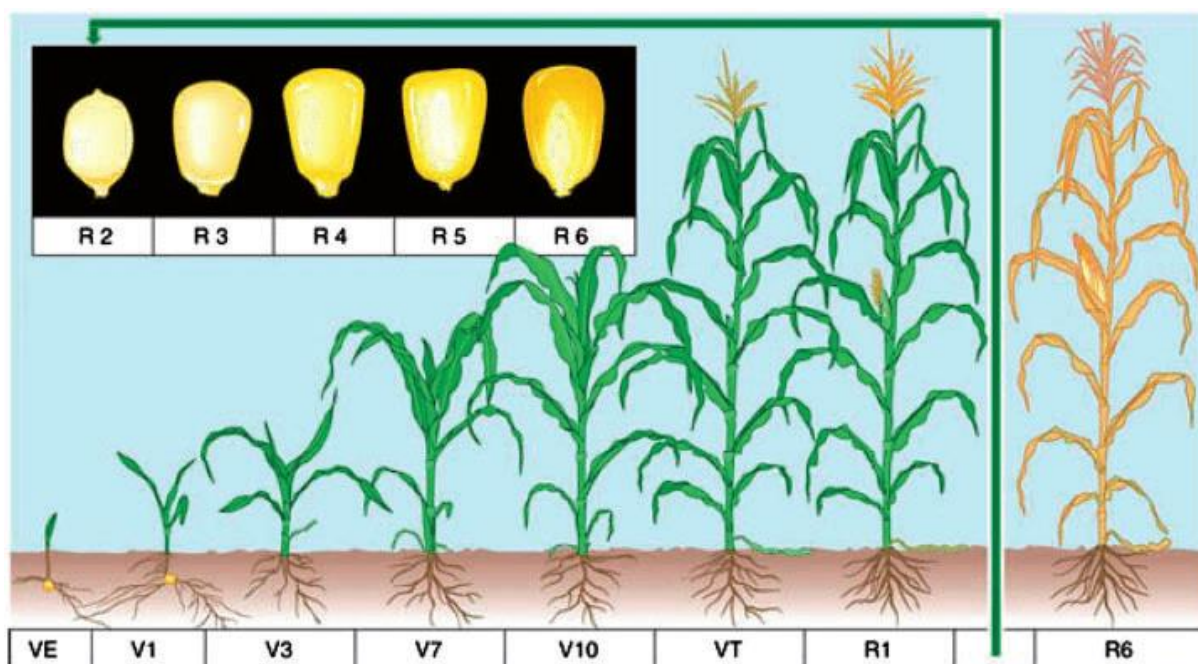


Figure 2 - Different growth stages of a maize plant, including the senescent stage (R6) (Zhao et al., 2011)

“Days to maturity” is often used to give an estimate for different maize varieties on when they reach physiological maturity. However, as definitions of physiological maturity can vary, and the duration a plant grows does not take into account the different climatic or soil conditions of individual parcels, “days to maturity” can only be considered as a rough indication (Hanway, 1963).

Kernel development is another way of assessing maturity: grain moisture, the position of the kernel milk line and the formation of kernel black layer have all been used to estimate corn maturity. Grain moisture is another indicator to determine harvest time. Maize is physiologically mature at about 30-35% moisture content in the grain. However, harvesting is most often delayed until the grain moisture content decreases to about 15–20%. Conventionally, a moisture content of about 13% is considered to be safe for storing the grain (Abendroth et al., 2009).

Observing the evolution of the kernel milk line is another method to assess the maturity of the plant. As the kernel develops from the “soft dough” to “no milk line visible” stages, predictions on the amount of dry matter and other yields are possible and are reliable (Tolera et al., 1998b). The formation of a black layer located at the base of the kernel where it is attached to the cob is also used as an indicator of maturity.

Afuakwa and Crookston (1984) have shown that both kernel milk line and kernel black layer can be used to reliably determine grain maturity, with kernel milk line being more useful for monitoring grain development before it reaches maturity. Both kernel milk line and black layer are visual methods and do not rely on any technical equipment, however, they are subject to individual judgement which might skew or falsify results. The grain moisture content cannot

be identified visually, and therefore relies on electronic equipment or time to dry the grain. Using grain moisture content to identify the optimal point to harvest requires periodic sampling and record keeping (Xu et al., 2019). All three methods are limited in that they cannot be applied to every single plot of an extended area of farmland and therefore do not satisfy the demands of modern farming.

Determining the optimal time of harvest is not easy for growers. Articles from *Farmersweekly* (Pringle, 2017) and from NC State (Heiniger, 2018) suggest that this is acknowledged in the farming community. Being able to assess correctly the plant water content, and thereby the maturity level to determine the optimal time of harvest for an extended area of farmland without the need of sampling would therefore be a significant improvement. Remote sensing could play an important role in achieving this.

2. Remote Sensing Data for Plant Water Content Crop monitoring

In the ever-evolving landscape of modern agriculture, satellite remote sensing has been emerging as a transformative force that is reshaping traditional farming paradigms. For over 40 years, the provision of timely, reliable, accurate and high-resolution satellite remote sensing imagery has facilitated a transition from parcel-level decision making towards precision agriculture (Haboudane et al., 2004; Mulla, 2013; Lamine et al., 2019). Being able to monitor vegetation and track crop conditions is expected to increase agricultural yields, improve current farming techniques, and help reduce the sector's carbon footprint (Prasad et al., 2006; Khabbazan et al., 2019).

2.1 Optical remote sensing

Optical remote sensing instruments – which use sensors to capture visible and near-infrared light – have been used extensively in the past decades to monitor crop yield at regional scales to support a broad range of uses. However, their resolution was not sufficient to capture the crop growth variability at a sufficient level of detail to be beneficial to farmers (Jongschaap and Schouten, 2005). Higher resolution optical satellites such as Sentinel-2 have since been launched and are today the most well established operational agricultural monitoring systems. These optical data have been used in several studies on the retrieval of biophysical characteristics in crops at various phenological stages, such as reported by Quarmby et al. (1993), Doraiswamy et al. (2004), Baret et al. (2007) and Bontemps et al. (2015). However, even when multiple high-resolution optical satellites are employed in a virtual constellation, cloud cover can block their sensors' view and hinder their use in implementing operational crop monitoring at field scales (Hosseini and McNairn, 2017).

2.2 Synthetic Aperture Radar (SAR)

Synthetic Aperture Radars (SAR) were developed as a means of overcoming the limitations of optical sensing equipment. SAR is a remote sensing imaging system that uses microwave electromagnetic emissions to create two-dimensional maps of radar brightness. Whereas optical sensing interprets photographs, SAR data responds to surface characteristics like structure and moisture. By recording and then combining these individual signals, a “synthetic aperture” is created, allowing the construction of a high-resolution complex image.

In contrast to optical sensors, active microwave sensors operate at longer wavelengths (centimetre to meter scale) which can penetrate clouds. SAR sensors are therefore able to acquire data even in the presence of cloud cover and are weather-independent. Their day-and-night imaging capability provides a continuous observational capacity.

Wavelength and frequency

When working with SAR, wavelength is an important feature to consider because it determines how the radar signal interacts with the surface and how far it can penetrate a medium. X-band, for example, having a small wavelength (~3 cm), cannot penetrate deeply into vegetation cover and will mostly interact with the leaves at the top. L-band on the other hand has a larger wavelength (~24 cm) and will penetrate deeper into the vegetation cover, interacting more with stems, trunks, and even the soil. This is illustrated in Figure 3.

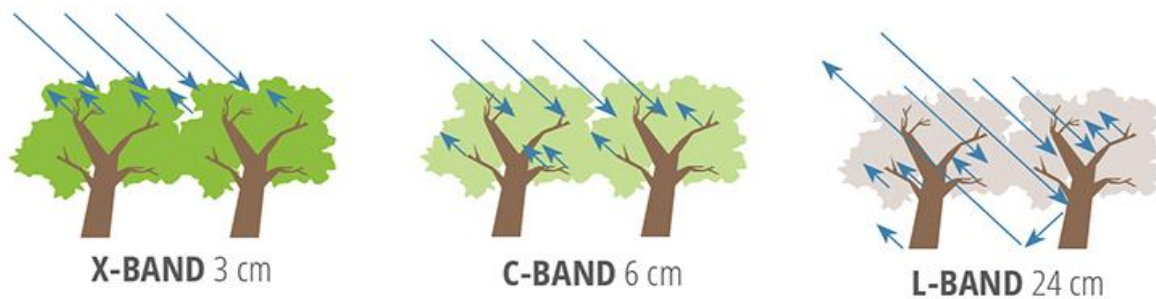


Figure 3 - Penetration of different SAR bands into vegetation cover (NASA SAR Handbook)

Different SAR bands have therefore specific applications based on their frequency and wavelength. The most used frequency bands for agriculture monitoring correspond to the X-, C- and L-band (Bouchat et al., 2022). Table 1 gives an overview of the bands most used in SAR when monitoring vegetation.

Table 1 - Overview of the most used frequencies and wavelengths in SAR and their application (NASA SAR Handbook)

Band	Frequency	Wavelength	Typical Application
X	8 - 12 GHz	2.3 - 3.8 cm	High-resolution SAR – urban monitoring, ice and snow; Little penetration into vegetation cover
C	4 - 8 GHz	3.8 - 7.5 cm	Reliable and frequently used – global mapping monitoring of areas; Low to moderate penetration
L	1 - 2 GHz	15 - 30 cm	Medium resolution SAR – geophysical monitoring, biomass and vegetation mapping; High penetration

The potential of low frequency (1-10 GHz) has long been recognized to provide recurrent and reliable observations essential for agricultural applications (Bush and Ulaby, 1978). In this frequency range, microwaves are sensitive to the water content and the structural changes in agricultural crops such as the size, shape, orientation, and roughness of the plants. The moisture content, roughness and texture of the underlying soil is also well detected (Khabbazan et al., 2019).

Polarisation and scattering mechanisms

Polarisation refers to the direction of travel of an electromagnetic wave. When a wave is transmitted horizontally, it oscillates on a plane parallel to the surface being imaged while a vertically transmitted wave oscillates on a plane that is at a right angle to the imaged surface.

SAR sensors typically use four combinations of polarisation: VV, VH, HV and HH, as shown in Table 2.

Table 2 - SAR polarisations (Introduction to SAR - HyP3, s.d.)

Polarisation Code	Transmit Signal Polarisation	Return Signal Polarisation
VV	Vertical	Vertical
VH	Vertical	Horizontal
HV	Horizontal	Vertical
HH	Horizontal	Horizontal

The initial letter represents the polarisation used to send the signal, and the second letter represents the polarisation of the measured return, as is depicted in Figure 4.

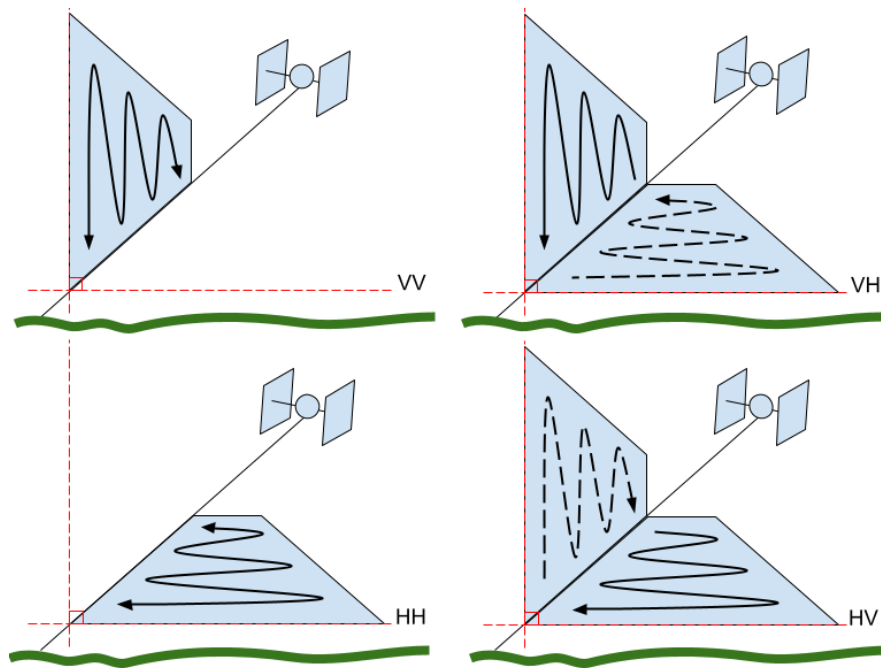


Figure 4 - SAR signals are transmitted and received either vertically (V) or horizontally (H). This gives the potential for four different polarization combinations (transmit listed first, receive second): VV, VH, HH, and HV (Introduction to SAR, s. d.)

The polarisation state of a backscattered wave from a natural surface can be linked to the geometrical characteristics (like shape, roughness and orientation) and the intrinsic properties of the scatterer (like humidity and moisture, salinity and dielectric properties) (*Echoes in space* – EO College, s. d.).

When an emitted microwave hits the surface of the Earth, it may reflect or scatter in various directions. If the radar pulse hits a smooth, flat surface (on the scale of the wavelength), most of the energy is scattered away in a specular direction. These areas will appear very dark in the radar image, such as water surfaces or roads. Three other types of scattering mechanisms exist: surface scattering, volume scattering, and double bounce, which are schematically represented in Figure . Surface scattering is mostly affected by the surface roughness and the wavelength size (the longer the wavelength, the smoother appears the surface for the sensor). Volume scattering happens when the radar pulse penetrates into a 3-dimensional body, such as a vegetated field, while double bounce occurs when the radar pulse hits two relatively smooth surfaces perpendicular to each other.

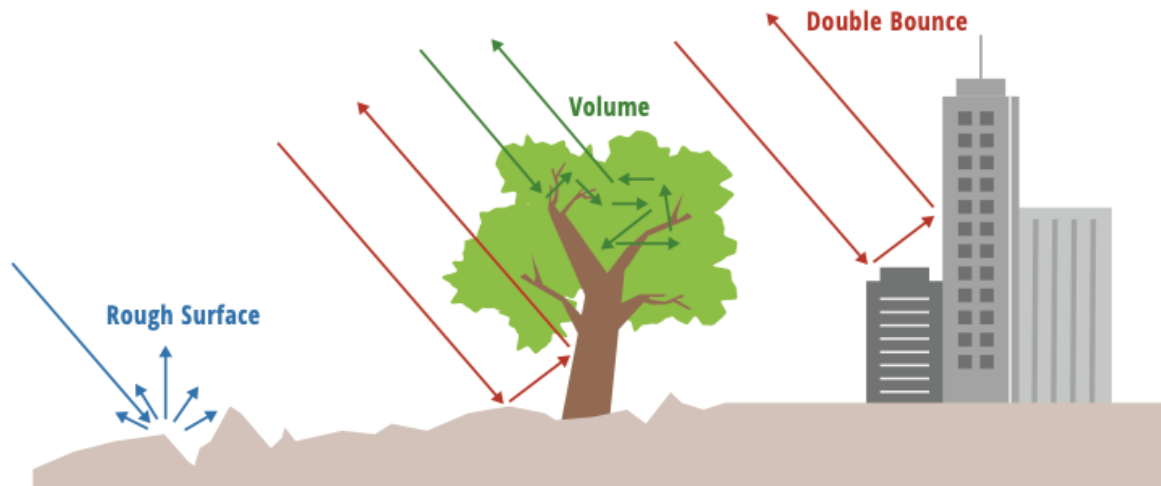


Figure 5 - SAR scattering mechanisms (Meyer, 2019)

By measuring the entire scattering matrix of a given object's backscatter coefficients, models can be developed to differentiate or identify the different scattering mechanisms that occur within a single pixel.

The strength of scattering depends on ground parameters (topography, surface roughness, object geometry and orientation, dielectric constant) and fixed sensor parameters (wavelength/frequency, polarisation, spatial resolution, incidence angle, orbit). These parameters will affect the backscattering coefficients. In natural land surfaces, scattering can be attributed to the soil, the vegetation or a combination of both. Typically, the backscattering signal is a result of various scattering mechanism (surface scattering, volume scattering, interaction scattering) but regardless of the scattering mechanism, the dielectric properties of the medium significantly influence the energy scattered back to the sensor, with an increase in the dielectric constant typically leading to an increase in backscatter. The most important dynamic driver for dielectric constant in soil and vegetation is the moisture content, thereby implying that backscatter generally increases with an increase in moisture content (Vreugdenhil et al., 2020).

Veloso et al. (2017) have mentioned that HH is more sensitive to surface scattering and HV to volume scattering, and VV a combination of the two. For this reason, HH backscatter is mostly used for ground parameters retrieval, and HV for crop parameters. Nonetheless, ground-based SAR measurements and theoretical modelling have shown that VV and VH can contain vegetation ground interaction (double-bounce effect) (Picard et al., 2003), with stem-ground interactions (double scattering) probably causing soil returns in VH polarisation (Brown et al., 2003). This assertion is backed by simulation results obtained from a second-order radiative transfer model (Brown et al., 2000; Wang et al., 2009; Liao et al., 2016)), revealing that double-bounce scattering exceeds the soil's direct backscatter. Veloso et al stated that the ratio VH/VV (Cross Ratio – CR) could be used to reduce this double-bounce effect. CR could indeed reduce

the acquisition system errors due to radiometric stability or external factors (such as variations in soil moisture) and emerges as a more consistent temporal indicator compared to VH or VV.

Vreugdenhil et al. (2020) compared CR from Sentinel-1 to vegetation dynamics as observed in microwave-based Vegetation Optical Depth (VOD) measurements from coarse-scale satellites (ASCAT) over Europe. They concluded that CR is useful for monitoring vegetation dynamics over crop- and grassland and a potential path to high-resolution VOD.

Speckle

The radar sensor's spatial resolution determines the smallest distance between two measurements that the sensor can distinguish, and it also influences the amount of speckle in the system. Speckle is a type of scattering that occurs when the sensor's spatial resolution is not high enough to differentiate individual scatterers. Unlike random noise, speckle can be replicated if the conditions are the same (Earth Science Data Systems, 2020). Although a homogeneous surface, such as a maize field, has an average backscatter coefficient, the coefficients can vary significantly between objects. Consequently, the radar image of the surface exhibits alternating bright and dark pixels, making a surface that appears uniform in optical images look grainy in radar images. This phenomenon creates challenges for photointerpretation and classification and must be addressed. To mitigate the impact of speckle, filters can be applied to the data (Lee et al., 1994; Lee et al., 2009). However, as trade-off, this may result a reduction in resolution or a blurring of features.

2.3 The use of SAR for vegetation monitoring and water content detection

Several papers have been published about crop classification methods thanks to the high temporal revisit of Sentinel-1, either solely with SAR data (Kussul et al., 2017; Torbick et al., 2017; Ndikumana et al., 2018; Bazzi et al., 2019), or combined with optical data (Bargiel, 2017). These papers offer understanding on the temporal patterns for a variety of crop types that are widely cultivated.

X-band

Although at first expected to have great potential (Blumberg, 2007), the use of X-band in vegetation and water content monitoring is limited because it only penetrates the upper part of the canopy.

Capodici et al. (2013) also admit that X-band is not the most suitable to model agricultural processes or monitor vegetation water content, but still achieve to correlate optical vegetation indices (Normalized Difference Vegetation Index – NDVI, and Leaf Area Index - LAI) with crossed Horizontal-Vertical (HV) backscattering from the COSMO SkyMed X-band SAR in

Italy. High correlation was found between NDVI estimated from σ_{HV}^0 and the one calculated from optical reflectance over corn fields ($R^2 \approx 0.98$).

C-band

Veloso et al. (2017) conducted a comparison between Sentinel-1 data and NDVI assessments derived from optical data in French fields. They also considered ground observations related to Green Area Index (GAI) and fresh biomass. They illustrated that the VH/VV ratio (CR) from Sentinel-1 could provide valuable insights into crop development by tracking the temporal changes in backscatter. Specifically for maize, they emphasised that there was a notable agreement between NDVI and CR with GAI and fresh biomass.

Vreugdenhil et al. (2018) have examined the potential of Sentinel-1 VV and VH backscatter and their ratio, the cross ratio (CR) to monitor crop conditions. They compared these microwave indices with vegetation variables of several crops measured in situ in Austria over two years. In their study, 6 parcels were studied with sizes varying between 2.3 and 9.4 ha, only two of which were used for growing maize. They investigated the backscatter patterns related to Plant Water Content (PWC), height, biomass, and Leaf Area Index (LAI). They demonstrated the applicability of Random Forest modelling to predict plant water content using Sentinel-1 data. The link between Sentinel-1 assessments and plant water content showed seasonal variations linked to the various plant development stages, with some periods showing backscatter changes mainly influenced by structural alterations. Their conclusion is that the CR dynamics reflect changes in the plant water content, illustrating the importance of analysing this polarisation ratio in the methodology of this report.

In the Netherlands, Khabbazan et al. (2019) obtained similar results showing that CR was sensitive to variations in crop water content and biomass. They also observed that CR was sensitive to structural changes in the vegetation. For the senescent stages of maize specifically, they saw a constant VH backscatter while the VV backscatter increased slightly resulting in an overall decrease of CR consistent with the conclusions of Vreugdenhil et al. (2018) that the CR dynamics reflect changes in vegetation water content.

In Catalonia, Spain, El Hajj et al. (2019) have retrieved VOD using a water cloud model from Sentinel-1 VV and VH backscatter observations in a test area of 50x50 km of agricultural fields which are not irrigated. Medium to good correlation coefficients for different crop types were found between Sentinel-1 VOD and NDVI from Sentinel-2 but no specific findings for maize are given in the study.

L-band

El Hajj et al. (2019) point out that due to the higher penetration depth of the L-band in a well-developed vegetation cover, it is more sensitive to the soil's moisture than the X- and C-bands and is therefore less pertinent to use in the case of vegetation monitoring.

Bouchat et al. (2022) also emphasise the importance of the frequency band in crop monitoring, because it influences the penetration depth of the signal into the canopy. In C-Band, a saturation phenomenon tends to occur for dense canopies. Backscatter in L-band, however, can penetrate a well-developed vegetation cover of wheat and grass to reach the ground beneath. L-band data may therefore also enable vegetation monitoring in maize fields into the later stages of plant development.

Chapter 3: Objectives

The objective of this master's thesis is to explore the potential of SAR data, specifically from TerraSAR-X and Sentinel-1 satellites, in assessing water content and biomass dynamics in senescent maize plants by comparing it with in situ-data from the Taiyuan Basin, China. To achieve this objective, the study focuses on two main axes: a comprehensive field data collection campaign, and the subsequent analysis of these data.

The first axis consists of collecting field data on senescent maize in the Taiyuan Basin, carried out jointly with our Chinese partners from the Taiyuan University of Technology during August and September 2023. This intensive field campaign involved gathering detailed measurements of various vegetation parameters in five different parcels of maize fields, including plant density, height, fresh and dry biomass, and water content. These in-situ data provide a robust foundation for understanding the physical and biological processes occurring in the maize plants during their senescent stage.

The second axis centres on the analysis of the collected data. A temporal analysis of biomass variables was performed to identify discernible patterns and trends in the maturation process of maize. These patterns are then integrated with concurrent SAR data to comprehensively understand the relationships between in-situ observations and remote sensing data. A linear and non-linear correlation analysis between SAR-derived metrics and field measurements were also conducted to quantify the relationships between remote sensing data and biophysical variables.

This study aims thereby to validate the capability of X- and C-bands to effectively monitor plant water content dynamics in senescent maize, particularly in semi-arid regions with small-scale field practices like the Taiyuan Basin.

Chapter 4: Materials and Methods

This chapter provides a detailed overview of the geographical and environmental context of the study area, focusing on the Taiyuan Basin in the Shanxi province of China. It outlines the site characteristics, field parcel selection criteria, and the unique challenges posed by the region's agricultural practices. Additionally, it introduces the main variables of interest, including fresh biomass, dry biomass, and water content, and describes the field campaign conducted between August and September 2023 to collect the necessary biophysiological data. The chapter concludes with a presentation of the satellite remote sensing data from TerraSAR-X and Sentinel-1 and details the subsequent analysis methods used to assess the water content in maize plants.

1. Region of Interest

1.1 Site description

To analyse the water content of maize plants, five parcels within the Taiyuan Basin in the Shanxi province of China were selected as features of interest. Situated in the North-Western part of China, the Taiyuan Basin is intersected by the Fen River, allowing for the establishment of irrigation infrastructure along its course. This abundant water resource supports agricultural activities and enhances the region's productivity.

Characterized by a semi-arid climate, the basin experiences temperatures ranging from a January mean of -7°C to a July mean of 24°C . Annual precipitation, primarily as rain, ranges from 400 to 650 mm, with 70 to 80 percent occurring from June to September, aligning with our field data collection. The influence of the southeastern monsoon is limited by the Taihang Mountains to the south of the basin (Boxer, 2024).

Winter irrigation predominates in the Taiyuan Basin region for maize production. A significant proportion of fields are irrigated during the winter and early spring seasons, despite being left as bare soil. This practice involves the application of large volumes of irrigation water to maintain adequate soil moisture levels for sowing crops at the onset of the growing season, which typically spans from May to October. The primary objective of this irrigation strategy is to ensure sufficient soil moisture availability for crop establishment, thereby mitigating potential water competition and preparing for potential drought conditions in the spring (Su et al., 2023).

1.2 Parcels characteristics

Five parcels have been selected by our Chinese partners during the intensive field campaign for this study. As shown in Figure 6, the selected parcels are distributed across the Taiyuan Basin to ensure representative sampling of the basin's area. The parcels are situated at considerable distances from each other, with the furthest distance being 77 km between Parcel 1 and Parcel 4. The parcels have a mean size of $20\,800\text{ m}^2$ (2.08 ha). Details of the studied parcels, including

their area, width, length, and coordinates, are provided in Table 3. This information was collected during the field campaign; additional data such as ownership, sowing and irrigation dates, and specific field practices were not available.

Table 3 - Parcel characteristics (area, width, length, and coordinates)

Parcel	Width [m]	Length [m]	Area [m ²]	Longitude (DMS)	Latitude (DMS)
1	150.43	180.68	24880.89	112° 42' 13"E	37° 38' 34"N
2	71.12	163.08	11379.12	112° 40' 04"E	37° 32' 01"N
3	100.70	310.99	31187.30	112° 07' 31"E	37° 24' 04"N
4	81.63	183.86	14495.08	111° 57' 04"E	37° 16' 41"N
5	106.27	212.93	22067.23	112° 22' 19"E	37° 19' 53"N

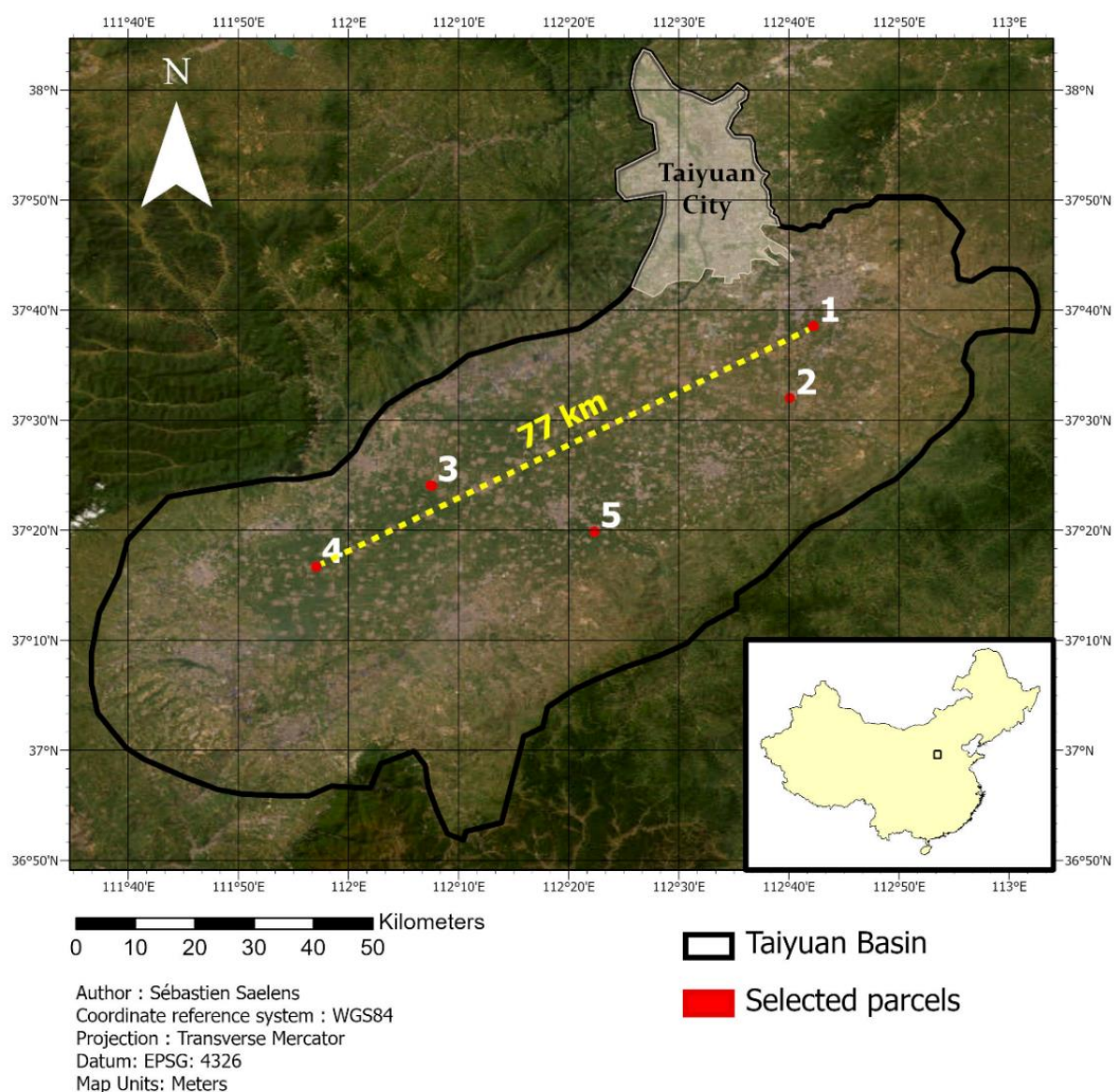


Figure 6 - Selected parcels within the Taiyuan Basin

A distinctive feature of cultivation practices in the region is the small size of fields, typically ranging between 5 and 20 meters in width. Consequently, the selected parcels encompass multiple fields, often held by different landowners and subject to diverse cultivation methods. To clarify, the term " Parcel " will be used to refer to groups of fields assembled into a single study area, corresponding to the different sampling zones. The term “field” refers to a plot of land belonging to the same farmer. A sketch of this situation can be found in Figure 7.

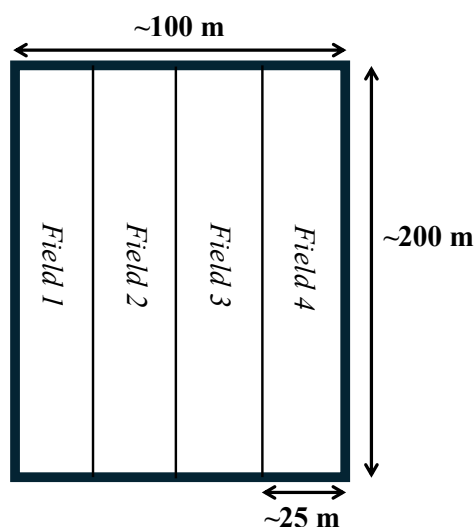


Figure 7 - Typical average parcel structure, containing multiple narrow fields managed by various owners leading to field heterogeneity. Number of fields and width may vary across parcels.

This diversity leads to variations in sowing dates, fertilizer and pesticide applications, and harvest schedules across various fields within the same parcel, contributing to significant parcel heterogeneity, as can be seen in Figure 8**Error! Reference source not found..** Field borders are sometimes cultivated with other vegetable crops such as cabbages, leeks or Chinese mustard, primarily intended for the farmers' personal use.



Figure 8 - Clear case of field heterogeneity, showing significant variations in plant height within the same parcel (Parcel 4)

The small average parcel size within the study area presents a significant challenge for assessing SAR sensor capabilities in detecting signal variations. With limited pixels per parcel, the inherent speckle effect in SAR images is accentuated, potentially obscuring variations in signal intensity. Moreover, noise pollution stemming from nearby features such as roads, power lines, or neighbouring crops distinct from maize, further complicates the analysis. Additionally, the heterogeneous nature of the fields within the parcels, characterised by variations in vegetation structure and water content, may introduce inconsistencies in the satellite signal, posing additional challenges to data interpretation.

To address these challenges, a trade-off was necessary: Either select single fields to minimize heterogeneity effects and accept the inherent increased speckle effect in SAR images, or group multiple fields within a single parcel to mitigate speckle effects but accept the increase in within-parcel heterogeneity. In this study, the latter option was chosen, as it allows for a more stable analysis despite the increased heterogeneity within the selected parcels.

2. Variables of Interest

The main variables of interest in this study are Fresh Biomass, Dry Biomass, and Water Content, which will be the focus of a statistical analysis to examine their relationship with TerraSAR-X and Sentinel-1 data.

2.1 Fresh biomass

Fresh biomass refers to the total mass of all above-ground plant material, including stems, leaves, and cobs, collected from the maize plants during field sampling. It is measured immediately after harvesting the plants and represents the current state of vegetation growth.

2.2 Dry biomass

Dry biomass represents the mass of plant material after removing all moisture through a drying process. It is measured by exposing the fresh biomass samples to controlled drying conditions until they reach a consistent weight, indicating complete moisture removal.

2.3 Water content

Water content (WC) is the amount of water present in the plant tissues, calculated as the difference between fresh biomass and dry biomass. It refers to the total mass of water in grams present in the fresh biomass samples. It can be converted to Area Water Content (AWC), measured in grams per square meter (g/m^2) using plant density measures taken on the parcels. This is the main variable that will affect the backscatter of the satellite signal.

3. Field Campaign

3.1 Field visit protocol

The five selected parcels were visited five times during a one-month period, starting on August 16, 2023, and ending on September 20, 2023, corresponding to the final growing stage of maize. Ideally, the biophysiological values should have been collected at the same time as SAR overpasses, with Sentinel-1 as the reference satellite. However, due to field constraints and rain hazards, biomass was collected within a maximum of 5 days preceding/following Sentinel-1 overpass because it is crucial to minimise any form of moisture, both in soil and on plants, during the collection of samples as excess moisture could distort the received signal, rendering it unrepresentative of plant water content. Sampling was initially planned on a weekly basis, but the last observation (week 5) was delayed due to field constraints outside our control. Details regarding the acquisition dates of in-situ measurements and radar backscatter acquisitions, along with the number of days between the in-situ measurements and the respective satellite passes, are provided in Table 4.

Table 4 - Acquisition dates for in-situ measurements and radar backscatter coefficients. #S1_Days and #TSX_Days indicate the number of days between the in-situ measurements and the respective satellite passage.

Parcel	Field date	S1 date	#S1_Days	TSX date	#TSX_Days
1	16-08-23	15-08-23	1	21-07-23	26
1	24-08-23	27-08-23	3	-	-
1	31-08-23	27-08-23	4	-	-
1	06-09-23	08-09-23	2	25-09-23	15
2	16-08-23	15-08-23	1	21-07-23	26
2	24-08-23	27-08-23	3	-	-
2	31-08-23	27-08-23	4	-	-
2	06-09-23	08-09-23	2	25-09-23	15
3	17-08-23	15-08-23	2	-	-
3	23-08-23	27-08-23	4	-	-
3	01-09-23	27-08-23	5	-	-
3	07-09-23	08-09-23	1	-	-
3	20-09-23	20-09-23	3	03-10-23	13
4	17-08-23	15-08-23	2	24-07-23	24
4	23-08-23	27-08-23	4	-	-
4	01-09-23	27-08-23	5	-	-
4	07-09-23	09-08-23	2	09-10-23	32
4	20-09-23	20-09-23	0	-	-
5	17-08-23	15-08-23	2	07-08-23	10
5	24-08-23	27-08-23	3	-	-
5	01-09-23	27-08-23	5	-	-
5	06-09-23	08-09-23	2	-	-

Upon arrival at the parcel, the first step was to assess the surrounding area to gauge weather conditions and their impact on the field and its vicinity. Another evaluation was carried out to assess heterogeneity within the parcel. Observations that could affect satellite signals, such as the presence of adventitious weeds or soil texture and roughness, were recorded and incorporated to support the results.

3.2 Field variables sampling

The vegetation sampling methods were derived from the ‘JOCAM Soil Moisture and Vegetation Measurement Guideline’ (Hosseini et al., 2018), and adapted to suit field conditions.

Planting Row Direction

The direction of planting was recorded in degrees using a compass and using magnetic North as a reference.

Plant Density

For row crops such as maize, both plant spacing (the distance between plants in each row) and row spacing (the distance between one plant row and the next) are needed to determine plant density. Plant density is needed to scale biomass samples to unit area (i.e., grams of plant per square meter).

To determine plant spacing, the mean number of plants along a transect of 10 meters was counted for five consecutive rows. Row spacing is determined by measuring the distance between rows, replicated for five rows. Plant density is then calculated using equation 1.

$$\text{Plant Density (PD)} = \frac{\text{Average Number of Plants in 5 Rows}}{(\text{Average Row Width over 5 rows}) \times 10\text{m}} \quad [1]$$

Plant density measures and row direction evaluation were only performed once for each parcel at the start of the field campaign.

3.3 Crop variables sampling

Crop height and biomass measurements were on the other hand performed during each parcel visit. The standard approach involved accessing the parcel from each of its four edges and venturing a minimum of 20 meters into the field to mitigate border effects. Samples from each side were selected to be as representative as possible of the parcel in terms of height, maturity, cob size, etc. The final sample was chosen towards the centre of the parcel, resulting in a total of five samples per parcel. This method was employed to ensure optimal representation of the parcel. A schematic illustration of the sampling approach can be found in Figure 9. All above-ground crop biomass, including out-of-earth roots, was collected using shears. The samples were then gathered at the edge of the field for subsequent measurements.

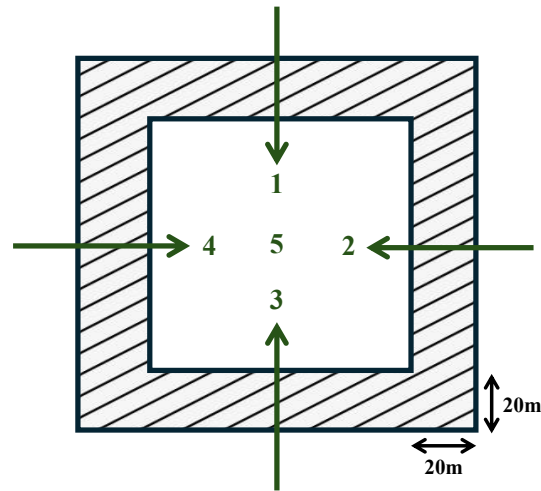


Figure 9 - Schematic sampling approach: The parcel is entered on all four sides and 4 samples are collected at least 20 meters inside the parcel. The last sample is collected in the centre of the parcel.

Crop Height

Crop height was assessed for each of the five harvested plants selected for biomass measurements. Typically, measures are conducted prior to destructive sampling; however, due to the considerable size of the plants - sometimes exceeding three meters in height as can be observed in Figure 10 - measurements were taken with the plants laid flat on the ground. The distance from the base of the stem to the highest point of the plants was determined using a tape measure, whilst ensuring that leaves were not extended during the measurement process. During each field visit, pictures were taken of every sample collection of every parcel for documentation and visual analysis purposes.



Figure 10 - Illustration of the considerable maize plant height, in comparison to me (1m84) (Parcel 2)

Fresh Biomass

To measure the fresh biomass, plants were first cut into smaller pieces and placed into mesh bags, as can be observed in Figure 11. Cobs were separated from the stems and leaves. Then, each bag was weighed with a scale accurate to within 0.01 grams. However, due to field conditions and practical considerations, the data was only read to the nearest 0.1 gram. The scale, positioned near the field, was carefully stabilized on the ground, which is not always flat. Special attention was paid to ensuring the weighing pan remained horizontal to ensure precise measurements.



Figure 11 - Jean Bouchat and our Chinese collaborators cutting maize samples using shears and placing them into mesh bags during a fieldwork session (Parcel 3)

Dry Biomass

After the collection of fresh biomass samples, each bag was transported to the Taiyuan University laboratory at the conclusion of each field trip and placed in a 60°C oven. The use of mesh bags allowed for increased airflow when placed in a drying room or oven, thereby expediting the drying process. Each bag was weighed daily, and the bags were removed from the oven only when the weight remained stable for three consecutive days. Stability in weight indicated that the sample was completely dry, as further evaporation of water had ceased.



Figure 12 - One of our Chinese partners measuring the dry biomass weight in the TYUT lab

Water Content

The Water Content (WC) was then determined based on the amount of water that evaporated during the drying process, using equation:

$$WC (g) = Fresh\ biomass (g) - Dry\ Biomass (g) \quad [2]$$

The Area Water Content (AWC), which indicated the amount of water in grams per square meter, was calculated using equation:

$$AWC (g/m^2) = \frac{WC (g)}{Number\ of\ plants\ collected} \times PD (plants\ per\ m^2) \quad [3]$$

The method for collecting the previously mentioned field variables is summarised in the following protocol:

1. Assess field conditions and heterogeneity for each parcel
2. Enter the parcel from each side and collect five representative samples
3. Measure the height of each sample and take a picture
4. Separate the cobs and place them in a mesh bag
5. Cut the plants into 10cm sections and put them into mesh bags
6. Weigh the fresh biomass bags
7. Repeat the sampling procedure for the next parcels
8. Once all parcels have been sampled, take the bags to the oven and leave them until their weight stabilizes
9. Weigh the dry biomass bags

4. SAR Data

4.1 X-band from TerraSAR-X

The first source of data used in this thesis comes from the TerraSAR-X satellite. TerraSAR-X is an earth observation satellite mission carried out under a public-private-partnership between the German Aerospace Center (DLR) and Airbus Defence and Space.

TerraSAR-X is a constellation of three satellites that collect X-band SAR imagery. The constellation includes TerraSAR-X and TanDEM-X, launched in June 2007 and June 2010 respectively. These two satellites operate in conjunction with a third, identical Spanish satellite called PAZ (Spanish for peace), which is owned and operated by Hisdesat.

Apart from vegetation monitoring, TerraSAR's X-band data is widely used for urban and infrastructure monitoring, as well as ice and snow monitoring, because the high-frequency radar band can detect objects with a diameter of up to a few centimetres (*TerraSAR-X - ESA Charter Mapper*, s. d.).

The available data, which consists of 23 images, were provided by Airbus. The imaging mode is called StripMap, up to 3m resolution with a scene size of approximately 30 km (width) × 50 km (length). The processing level of the imagery used is Enhanced Ellipsoid Corrected (EEC). This is the highest level of preprocessing, which allows a quick and highly accurate overlay with other spatial data. An example of an image used for this research is provided Figure 13.

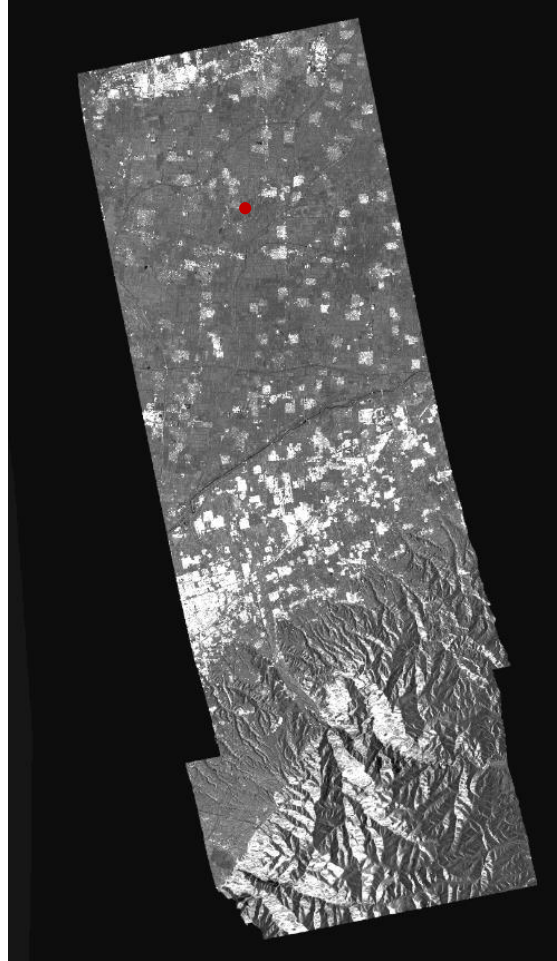


Figure 13 - TSX image taken on 10 October; comprising Parcel 4 (red dot)

The data were processed using the Sentinel Application Platform (*SNAP - ESA Sentinel Application Platform v9.0.0*, 2022) provided by the European Space Agency. Radiometric calibration was performed so that the pixel values of the SAR images truly represent the radar backscatter of the reflecting surface which in any case is also necessary for the comparison with Sentinel-1 images. Zonal statistics were then calculated on each of the five parcels, to obtain backscattering coefficients σ^0 for each parcel polygon. After processing, only nine data points were extracted from the 23 provided imagery. Essentially, the parcels were completely outside the boundaries of most TerraSAR-X images, leading to an inability to extract meaningful backscatter data from those images. Such an image can be seen at Figure 14. Consequently, only the images that fully encompassed the parcels could be used, resulting in the limited number of data points available. These data points are displayed in Table 5.

Table 5 - TerraSAR-X available datapoint, including parcel number, date and time of acquisition and various backscatter coefficients

Parcel	TSX datetime	VH: mean	VH: std	VH: median	VH: min	VH: max	VV: mean	VV: std	VV: median	VV: min	VV: max
1	2023-07-21 22:45:17	0.15965	0.12039	0.12504	0.00489	0.81562	0.14453	0.10866	0.11692	0.00466	0.86683
1	2023-09-25 22:45:21	0.31200	0.25581	0.24178	0.01514	2.63092	0.25448	0.20220	0.19676	0.00950	1.82776
2	2023-07-21 22:45:17	0.18846	0.15376	0.14294	0.00623	1.19458	0.14015	0.10202	0.11621	0.00458	0.74696
2	2023-09-25 22:45:21	0.32857	0.25833	0.24970	0.01013	2.02358	0.26196	0.21361	0.19863	0.00703	1.35893
3	2023-10-03 10:30:34	0.04704	0.03139	0.04013	0.00037	0.35906	0.05112	0.03392	0.04341	0.00009	0.33836
4	2023-07-24 10:21:55	0.19224	0.14399	0.15448	0.01398	1.42581	0.14897	0.10598	0.12354	0.00537	0.73258
4	2023-10-09 10:21:54	0.14328	0.10355	0.11858	0.00195	0.75066	0.11693	0.07369	0.09793	0.00695	0.49701
4	2023-10-09 10:22:00	0.14754	0.10540	0.11967	0.00314	0.69930	0.12149	0.07670	0.10243	0.00753	0.50964
5	2023-08-07 22:36:54	0.15350	0.11733	0.12440	0.00030	1.05324	0.12973	0.10154	0.10195	0.00023	0.82001

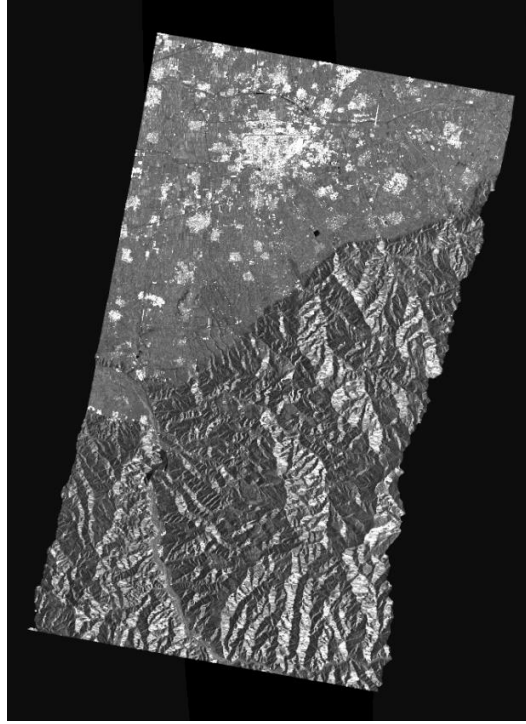


Figure 14 - Example of TSX image not including any of the 5 parcels (21 July)

4.2 C-band from Sentinel-1

Sentinel-1 is a component of the Copernicus programme of Europe, currently comprising two orbiting satellites, Sentinel-1A and Sentinel-1B, which were launched in April 2014 and 2016, respectively. These Sentinel-1 satellites are equipped with Synthetic Aperture Radars (SAR) that deliver backscatter at the C-band frequency (5.405 GHz). Sentinel-1 provides co- and cross-polarised backscatter data over a swath of 250 km with a spatial resolution of 20 m in a single look. A single Sentinel-1 satellite revisits a location every 12 days.

C-band SAR data acquired by the S1 constellation in interferometric wide (IW) swath mode are received in level-1 single-look complex (SLC) format. For this research, the available satellite images (205) range from 30 December 2022 to 19 November 2023.

These images were pre-processed using ESA's Sentinel-1 toolbox (S1TBX) with the Sentinel Application Platform (*SNAP - ESA Sentinel Application Platform v9.0.0*, 2022). Backscattering coefficients were obtained through several steps:

1. TOPSAR split
2. Orbit file application
3. Thermal noise removal
4. Calibration to σ^0
5. TOPSAR deburst
6. TOPSAR merge

These preprocessing steps were realised by the Earth and Life Institute – Environmental Sciences (ELIE) at UCLouvain, unlike the TerraSAR-X images, which were processed and calibrated by myself. Consequently, the work presented in this thesis focused solely on the derived dataset from Sentinel-1, without directly handling the raw images.

The derived dataset consisted of 125 data points over the entire year, in contrast to the available TSX dataset that ranges from 21 July to 9 October. A time series illustrating the evolution of the cross-ratio for all parcels during the whole growing season is provided in Figure 15.



Figure 15 - Temporal evolution of the cross ratio (CR) for the year 2023 for all maize parcels

The received dataset consists of a time series over the whole year. However, only 22 datapoints were selected to align with the timing of in-situ biomass measurements, as shown in Table 6. For all parcels, the data points from August 27 are between two in-situ measurements and have therefore been duplicated to match the nearest in-situ measurements.

Table 6 - Sentinel-1 selected datapoints, including parcel number, date and time of acquisition and various backscatter coefficients

Parcel	S1 datetime	VH: mean	VH: std	VH: median	VH: min	VH: max	VV: mean	VV: std	VV: median	VV: min	VV: max
1	2023-08-15 10:29:47	0.01796	0.01570	0.01304	0.00010	0.07489	0.07842	0.06028	0.06097	0.00385	0.31423
1	2023-08-27 10:29:47	0.02310	0.01942	0.01888	0.00015	0.13125	0.08932	0.08477	0.06122	0.00134	0.57029
1	2023-08-27 10:29:47	0.02310	0.01942	0.01888	0.00015	0.13125	0.08932	0.08477	0.06122	0.00134	0.57029
1	2023-09-08 10:29:48	0.01554	0.01436	0.01138	0.00002	0.08386	0.07256	0.05959	0.05419	0.00184	0.38119
2	2023-08-15 10:29:47	0.01501	0.01544	0.01071	0.00002	0.07519	0.08457	0.08219	0.05966	0.00428	0.44528
2	2023-08-27 10:29:47	0.02376	0.01845	0.01856	0.00054	0.08895	0.11520	0.11686	0.08398	0.00622	0.72418
2	2023-08-27 10:29:47	0.02376	0.01845	0.01856	0.00054	0.08895	0.11520	0.11686	0.08398	0.00622	0.72418
2	2023-09-08 10:29:48	0.01675	0.01406	0.01425	0.00007	0.07642	0.07934	0.06617	0.06261	0.00293	0.34534
3	2023-08-15 10:29:47	0.01834	0.01565	0.01385	0.00023	0.08758	0.07785	0.06514	0.06227	0.00190	0.44965
3	2023-08-27 10:29:47	0.02478	0.02168	0.01800	0.00001	0.13307	0.08286	0.08173	0.05916	0.00420	0.78064
3	2023-08-27 10:29:47	0.02478	0.02168	0.01800	0.00001	0.13307	0.08286	0.08173	0.05916	0.00420	0.78064
3	2023-09-08 10:29:48	0.02170	0.02476	0.01617	0.00053	0.26592	0.07783	0.06321	0.06016	0.00370	0.47695
3	2023-09-20 10:29:48	0.01803	0.01384	0.01440	0.00025	0.07824	0.06518	0.04752	0.05295	0.00129	0.34955
4	2023-08-15 10:29:47	0.02252	0.01485	0.02092	0.00081	0.08141	0.05624	0.04461	0.04681	0.00313	0.25198
4	2023-08-27 10:29:47	0.02292	0.01813	0.01792	0.00029	0.08041	0.09619	0.06648	0.07295	0.00658	0.31886
4	2023-08-27 10:29:47	0.02292	0.01813	0.01792	0.00029	0.08041	0.09619	0.06648	0.07295	0.00658	0.31886
4	2023-09-08 10:29:48	0.01655	0.01947	0.01072	0.00001	0.12511	0.08111	0.06067	0.06500	0.00475	0.30613
4	2023-09-20 10:29:48	0.01784	0.01790	0.01338	0.00013	0.11391	0.08142	0.05589	0.06461	0.00480	0.30000
5	2023-08-15 10:29:47	0.02026	0.02233	0.01341	0.00001	0.15324	0.08195	0.06643	0.06470	0.00255	0.38567
5	2023-08-27 10:29:47	0.02508	0.02682	0.01628	0.00022	0.17354	0.11040	0.12006	0.08246	0.00480	1.03099
5	2023-08-27 10:29:47	0.02508	0.02682	0.01628	0.00022	0.17354	0.11040	0.12006	0.08246	0.00480	1.03099
5	2023-09-08 10:29:48	0.02211	0.02383	0.01448	0.00001	0.17959	0.11124	0.14941	0.07993	0.00293	1.81386

4.3 Data format

The satellite data is presented in XLSX format and includes information about each satellite observation, such as the date and time of observation, platform, orbit, count, and pass (for S1). Additionally, it contains the mean, median, maximum, minimum, and standard deviation for the VV and VH polarisations. Several radar indices have been calculated and added to the dataset.

The mean backscatter values have been converted to decibels for both polarisations using the equation:

$$Mean \sigma^0 [dB] = 10 \times \log_{10}(Mean \sigma^0) \quad [4]$$

The VH/VV ratio, also known as the cross-ratio (CR) has been calculated using the equation:

$$CR [dB] = \frac{\sigma_{VH:mean}^0}{\sigma_{VV:mean}^0} = VH:mean \text{ dB} - VV:mean \text{ dB} \quad [5]$$

The Dual-Polarisation Radar Vegetation Index (dpRVI), an index commonly used to compare vegetation properties based on radar data limited to two polarisations (VH and VV) (Kim and Van Zyl, 2009; Mandal et al. 2020), has also been calculated using the equation:

$$dpRVI = 4\sigma_{VH}^0 / (\sigma_{VV}^0 + \sigma_{VH}^0) \quad [6]$$

Overviews of the data formats for TerraSAR-X and Sentinel-1 are shown in Table 7 and Table 8. These tables serve as a visualisation of the data format used to produce the results of this research.

Table 7 - TSX dataset with additional radar indexes (VH and VV: mean in decibels, CR and dpRVI), including parcel number, date and time of acquisition, the platform, the orbit number, and TSX count.

Parcel	TSX datetime	Platform	Orbit	TSX count	VV: mean dB	VH: mean dB	CR	dpRVI
1	2023-07-21 22:45:17	TDX1	101	3289	-8.4005	-7.9682	0.4323	2.0995
1	2023-09-25 22:45:21	TDX1	101	3289	-5.9434	-5.0584	0.8850	2.2031
2	2023-07-21 22:45:17	TDX1	101	1500	-8.5342	-7.2478	1.2864	2.2940
2	2023-09-25 22:45:21	TDX1	101	1500	-5.8177	-4.8337	0.9839	2.2256
3	2023-10-03 10:30:34	TDX1	101	4123	-12.9138	-13.2753	-0.3615	1.9168
4	2023-07-24 10:21:55	TDX1	101	1916	-8.2691	-7.1615	1.1076	2.2537
4	2023-10-09 10:21:54	TDX1	101	1921	-9.3206	-8.4380	0.8826	2.2025
4	2023-10-09 10:22:00	TDX1	101	1916	-9.1545	-8.3109	0.8435	2.1936
5	2023-08-07 22:36:54	TSX1	101	2919	-8.8696	-8.1389	0.7306	2.1678

Table 8 - S1 dataset with additional radar indexes (VH and VV: mean in decibels, CR and dpRVI), including parcel number, date and time of acquisition, the platform, the orbit number, and S1 count.

Parcel	S1 datetime	Platform	Orbit	S1 count	VV: mean dB	VH: mean dB	CR	dpRVI
1	2023-08-15 10:29:47	S1A	113	242	-11.0559	-17.4569	-6.4010	0.7454
1	2023-08-27 10:29:47	S1A	113	252	-10.4904	-16.3648	-5.8745	0.8218
1	2023-08-27 10:29:47	S1A	113	252	-10.4904	-16.3648	-5.8745	0.8218
1	2023-09-08 10:29:48	S1A	113	246	-11.3929	-18.0848	-6.6919	0.7056
2	2023-08-15 10:29:47	S1A	113	111	-10.7278	-18.2353	-7.5074	0.6030
2	2023-08-27 10:29:47	S1A	113	114	-9.3855	-16.2422	-6.8567	0.6839
2	2023-08-27 10:29:47	S1A	113	114	-9.3855	-16.2422	-6.8567	0.6839
2	2023-09-08 10:29:48	S1A	113	110	-11.0050	-17.7610	-6.7560	0.6971
3	2023-08-15 10:29:47	S1A	113	312	-11.0872	-17.3649	-6.2777	0.7628
3	2023-08-27 10:29:47	S1A	113	312	-10.8167	-16.0583	-5.2416	0.9210
3	2023-08-27 10:29:47	S1A	113	312	-10.8167	-16.0583	-5.2416	0.9210
3	2023-09-08 10:29:48	S1A	113	310	-11.0886	-16.6345	-5.5459	0.8722
3	2023-09-20 10:29:48	S1A	113	312	-11.8587	-17.4404	-5.5818	0.8666
4	2023-08-15 10:29:47	S1A	113	147	-12.4992	-16.4736	-3.9743	1.1438
4	2023-08-27 10:29:47	S1A	113	145	-10.1686	-16.3976	-6.2290	0.7697
4	2023-08-27 10:29:47	S1A	113	145	-10.1686	-16.3976	-6.2290	0.7697
4	2023-09-08 10:29:48	S1A	113	142	-10.9092	-17.8114	-6.9023	0.6779
4	2023-09-20 10:29:48	S1A	113	144	-10.8929	-17.4866	-6.5937	0.7189
5	2023-08-15 10:29:47	S1A	113	220	-10.8647	-16.9332	-6.0685	0.7930
5	2023-08-27 10:29:47	S1A	113	215	-9.5703	-16.0071	-6.4368	0.7404
5	2023-08-27 10:29:47	S1A	113	215	-9.5703	-16.0071	-6.4368	0.7404
5	2023-09-08 10:29:48	S1A	113	221	-9.5375	-16.5542	-7.0167	0.6632

5. Data Analysis and Processing

Once all the variables are collected, they are analysed and correlated using the Python programming language in various Jupyter Notebooks. The assistance of ChatGPT is used to refine and debug the code.

First, an exploratory analysis of the in-situ data is carried out to identify trends observed during the final maturation stages. It is expected that biomass will decrease over time, while dry matter will increase as the plant undergoes dehydration in the senescence stage. Therefore, a positive correlation is anticipated between biomass and water content. Time series of fresh and dry biomass are performed to observe these trends.

Similar univariate analyses are performed for the satellite backscatter coefficients, specifically for VH and VV mean values and the cross-ratio. These remote sensing data are also compared to the duration of the field campaign, although for TerraSAR-X, the data points are not concurrent with the in-situ data.

Next, each in-situ variable is correlated with the satellite backscatter coefficients (in the VH and VV polarisations). This correlation analysis is used to calculate the correlation coefficient for each variable. The highest correlation is anticipated between the area water content (measured in g/m^2) and the backscatter coefficients. Biomass is also anticipated to exhibit a strong correlation.

In addition to linear univariate correlation analysis, non-linear multivariate correlation trends are explored using machine learning techniques. Random Forest Regressors are employed using the Scikit-learn library in Python to examine potential non-linear relationships between the in-situ variables and satellite backscatter coefficients. This multivariate approach allows for the detection of complex non-linear patterns that are not captured by traditional correlation methods. Furthermore, leave-one-out cross-validations are conducted to assess the robustness and generalisation of the observed correlations to unseen data.

Chapter 5: Results and Interpretations

This chapter presents the raw data collected in the field using the methodology explained in the previous section. It then presents the results of linking these field data to each other and to the available SAR data. The section concludes by the results of a correlation analysis between the various variables of interest.

1. Field Campaign

The following tables summarise the measurements obtained during the field campaign, including the date and time of sample collection, and the sum of the five samples collected (Fresh Biomass, in grams). The Dry Biomass (in grams) was measured in the lab after drying out the samples, while other variables were derived afterward. These tables serve as foundational datasets for the subsequent analysis and interpretation of biomass and water content evolution presented in the Results and Discussion chapters. Additional comments are also made whenever external events may affect the results, such as irrigation, rain, fog, etc...

1.1 First Field Visit (16-17 August)

Parcels 1 to 4 were very dry during the initial field visit, while the soil of Parcel 5 was notably humid, likely due to irrigation two to three days prior (according to passers-by).

Table 9 - Variables collected during the first field visit (16-17 August)

Parcel	Date	Time	Fresh Biomass [g]	Dry Biomass [g]	WC [g]	Area WC [g/m ²]	WC [%]	Dry Matter [%]
1	16-08-23	9.20	5417.0	1559.3	3857.7	2193.0	71.2	28.8
2	16-08-23	16.30	5052.0	1455.6	3596.4	2036.0	71.2	28.8
3	17-08-23	10.00	6141.0	1450.4	4690.7	2995.2	76.4	23.6
4	17-08-23	15.15	7187.0	1636.4	5550.6	2614.7	77.2	22.8
5	17-08-23	18.45	6437.0	1760.3	4676.8	2651.1	72.7	27.3

1.2 Second Field Visit (23-24 August)

Light rain was observed on 23 August during sampling of Parcel 4, potentially impacting water content measurements for this parcel. The rain may have continued overnight, with fog and dew present during the sampling process on the morning of the 24th. The first rays of sun appeared at 08:50, prompting a delay of around two hours in sample bagging for Parcel 1 to ensure maximum drying in the sunlight. Samples were still slightly humid.

Concerning Parcel 5, the soil was drier than the previous week, likely indicating the suspension of irrigation. Despite this, the parcel remained wetter than others visited on that day.

Table 10 - Variables collected during the second field visit (23-24 August)

Parcel	Date	Time	Fresh Biomass [g]	Dry Biomass [g]	WC [g]	Area WC [g/m ²]	WC [%]	Dry Matter [%]
1	24-08-23	8.30	5132.0	1834.4	3297.6	1874.6	64.3	35.7
2	24-08-23	12.50	5463.0	1787.9	3675.1	2080.5	67.3	32.7
3	23-08-23	9.20	5348.0	1399.2	3948.8	2521.5	73.8	26.2
4	23-08-23	14.30	7036.0	1701.2	5334.8	2513.0	75.8	24.2
5	24-08-23	16.00	5601.0	1935.9	3665.1	2077.6	65.4	34.6

1.3 Third Field Visit (31 August – 1 September)

On 29 August it had been raining heavily. Despite the wet soil, the plants were perfectly dry to the touch on 31 August. In the case of Parcel 3, six samples were collected to better assess field heterogeneity. To maintain consistency with other biomass measurements where five samples were collected, the weight was normalised.

Significant lodging spots, where plants had fallen over because of animals or heavy rain, were observed in the centre of Parcels 2 and 5. The cobs of Parcel 5 were markedly dry, indicating that harvest was imminent. A small section of this parcel (approximately 3 metres in width) had already been harvested since the last visit.

Table 11 - Variables collected during the third field visit (31 August - 1 September)

Parcel	Date	Time	Fresh Biomass [g]	Dry Biomass [g]	WC [g]	Area WC [g/m ²]	WC [%]	Dry Matter [%]
1	31-08-23	12.45	4432.0	1897.1	2535.0	1441.0	57.2	42.8
2	31-08-23	16.15	5334.0	1841.9	3492.1	1976.9	65.5	34.5
3	01-09-23	8.15	5175.0	1500.4	3674.6	2346.4	71.0	29.0
4	01-09-23	11.30	5952.0	1634.0	4318.0	2034.0	72.5	27.5
5	01-09-23	15.30	5234.0	1959.0	3275.0	1856.5	62.6	37.4

1.4 Fourth Field Visit (6-7 September)

The soil was notably drier, suggesting that in all likelihood there had been no rainfall since the last visit. The maize in all parcels appeared to have reached maturity, except for Parcel 4, which seemed to be the least mature of all sites. Maize in this parcel remained very green and not yet dry. Additionally, in this parcel, numerous small insects, identified as aphids, were observed in the upper part of the plants.

Table 12 - Variables collected during the fourth field visit (6-7 September)

Parcel	Date	Time	Fresh Biomass [g]	Dry Biomass [g]	WC [g]	Area WC [g/m ²]	WC [%]	Dry Matter [%]
1	06-09-23	7.45	4462.0	1954.7	2507.4	1425.3	56.2	43.8
2	06-09-23	11.40	4689.0	1927.9	2761.1	1563.1	58.9	41.1
3	07-09-23	8.40	4411.7	1529.0	2882.7	1840.7	65.3	34.7
4	07-09-23	12.40	5821.0	2083.1	3737.9	1760.8	64.2	35.8
5	06-09-23	15.30	4933.0	2038.3	2894.7	1640.9	58.7	41.3

1.5 Fifth Field Visit (20 September)

The final field visit was delayed by one week because of prior commitments of our Chinese collaborators. Two weeks after the last visit, all parcels, except Parcel 4, had been completely

or partially harvested. Sampling measures were not performed in parcels where the harvest exceeded 50% of the total surface area (Parcel 1, 2 and 5).

Table 13 - Variables collected during the fifth field visit (20 September)

Parcel	Date	Time	Fresh Biomass [g]	Dry Biomass [g]	WC [g]	Area WC [g/m ²]	WC [%]	Dry Matter [%]
3	20-09-23	16.00	4826.0	1887.9	2938.1	1670.2	60.9	39.1
4	20-09-23	13.45	5306.0	1896.4	3409.6	1938.2	64.3	35.7

2. Temporal Evolution of Variables of Interest

This section presents graphs illustrating the temporal evolution of the variables of interest throughout the field campaign. It is important to note that while the data generally exhibit the expected trends, small incoherences were to be expected and were observed. These discrepancies can be attributed to various factors inherent to the field campaign process, including sampling selection not fully representative of the parcel, measurement errors, inconsistencies in drying times, or possible encoding errors. Despite these inconsistencies, the data provide valuable insights into the dynamics of the studied variables.

2.1 Fresh biomass

Figure 16 shows the temporal evolution of the fresh biomass in grams for each parcel during the field campaign (from 16/08 to 20/09). Fresh biomass represents the sum of the biomass of all five samples collected in each parcel on a particular field visit. These trends highlight the progression of the crop maturity.

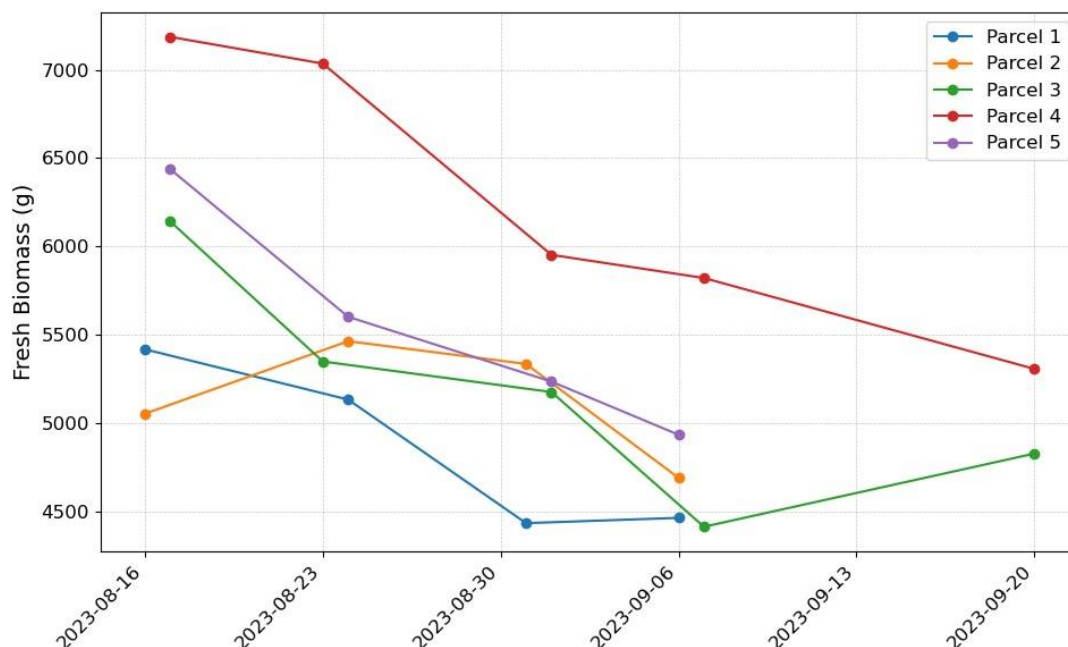


Figure 16 - Temporal evolution of the fresh biomass in grams for the five parcels (16/08 – 20/09)

As expected, fresh biomass exhibits a declining trend over time, which aligns with the expected drying process of the maize crop in the senescent stage. Parcel 4 is with an initial mass of 7187 g the parcel with the highest fresh biomass, which is consistent with field observations. It was

noted that the plants in this parcel were greener and at a less advanced stage compared to other parcels. The in-situ data generally appears accurate, although small deviations are noticeable. The initial fresh biomass sample for Parcel 2 appears unusually low, while Parcel 3 demonstrates a resurgence in biomass during the final sampling stage. As explained at the beginning of this section, these inconsistencies are probably linked to the various errors inherent to the field campaign process.

2.2 Dry Biomass

Figure 17 illustrates the temporal evolution of the dry biomass in grams for each parcel over the field campaign period. Dry biomass is calculated as the total mass of the five samples collected on each parcel visit, after drying in the oven.

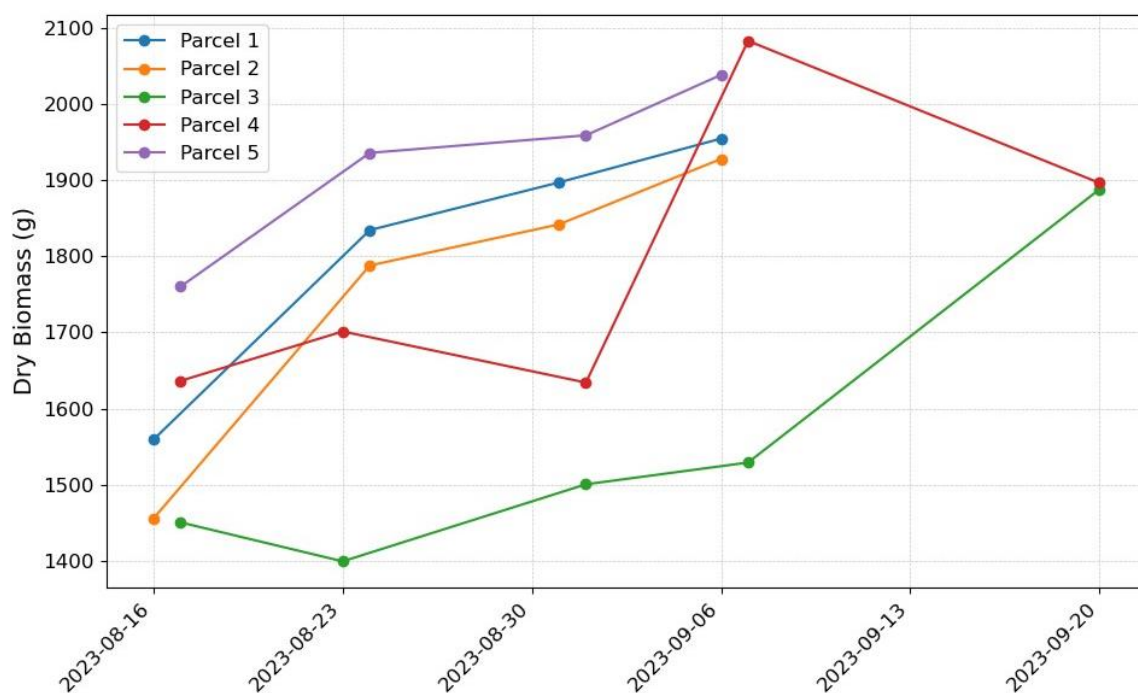


Figure 17 - Temporal evolution of the dry biomass in grams for the five parcels (16/08 – 20/09)

Similarly, dry biomass is increasing as expected while the maize undergoes gradual desiccation. As sharp increase is observed between the initial and subsequent field visits for Parcel 1, 2, and 5, contrasting with more modest changes in Parcel 4 and even a decrease in Parcel 3. Additionally, the data for Parcel 4 exhibit less smoothness compared to the other parcels, displaying a somewhat erratic pattern instead of a consistent rising tendency observed in other parcels.

By using the data of the fresh and the dry biomass, it is possible to calculate the dry biomass ratio evolution, shown in Figure 18, using the formula:

$$\text{Dry Biomass Ratio [\%]} = \frac{\text{Dry Biomass [g]}}{\text{Fresh Biomass [g]}} \times 100 \quad [7]$$

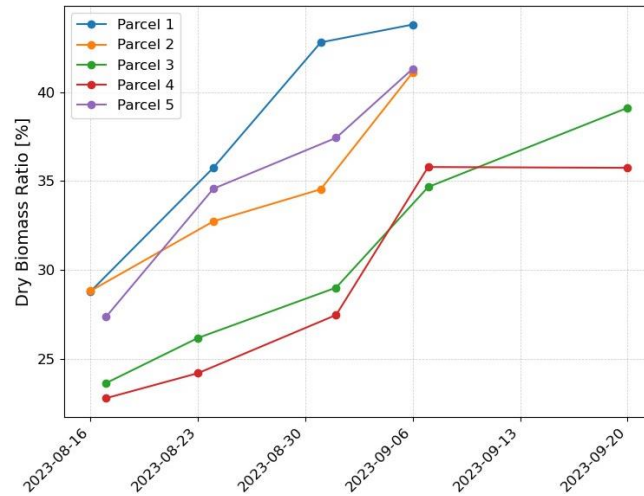


Figure 18 - Dry Biomass Ratio in percentage for the five parcels (16/08 – 09/20)

It can be observed that the tendency is increasing for all parcels. Parcel 4 is however showing a stabilisation in the dry biomass ratio for the last field visit.

2.3 Water Content

Figure 19 presents the temporal evolution of water content in grams for each parcel during the field campaign. The water content was derived by subtracting the dry biomass from the fresh biomass (see equation 2).

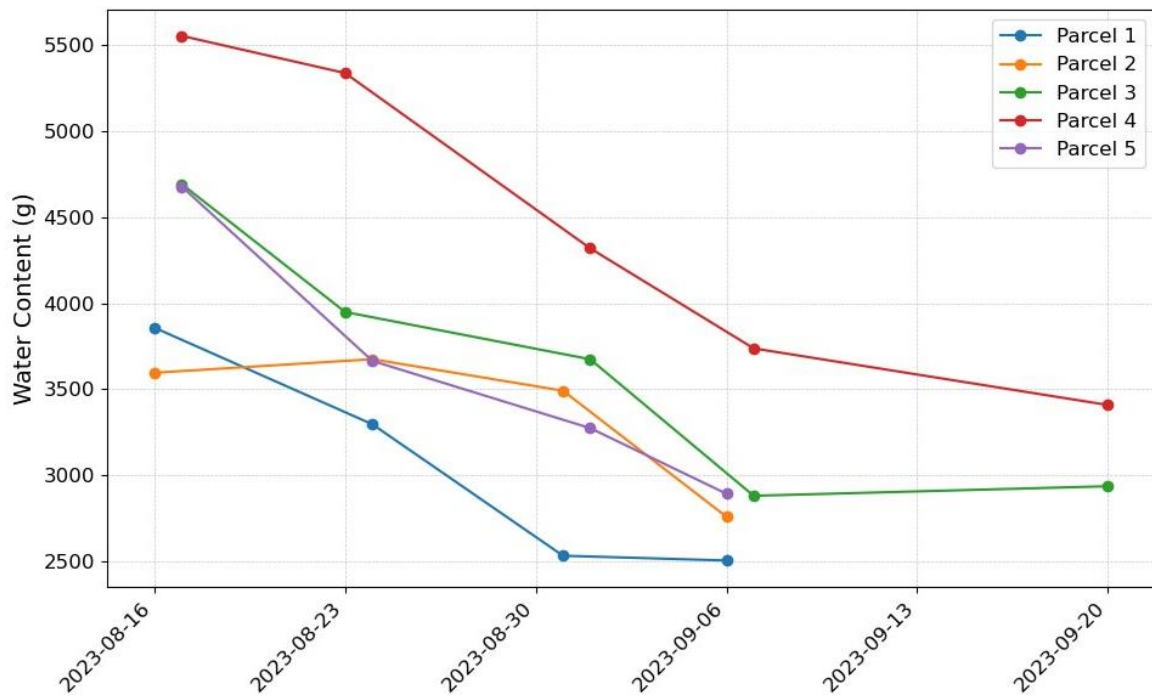


Figure 19 - Temporal evolution of the water content for the five parcels (16/08 - 20/09)

The Water Content Ratio in percentage that is shown on Figure 20 can be calculated using the formula:

$$\text{Water Content Ratio [\%]} = \frac{\text{Water Content [g]}}{\text{Fresh Biomass [g]}} \times 100 \quad [8]$$

$$\text{Or simply: Water Content Ratio [\%]} = 1 - \text{Dry Biomass Ratio [\%]} \quad [9]$$

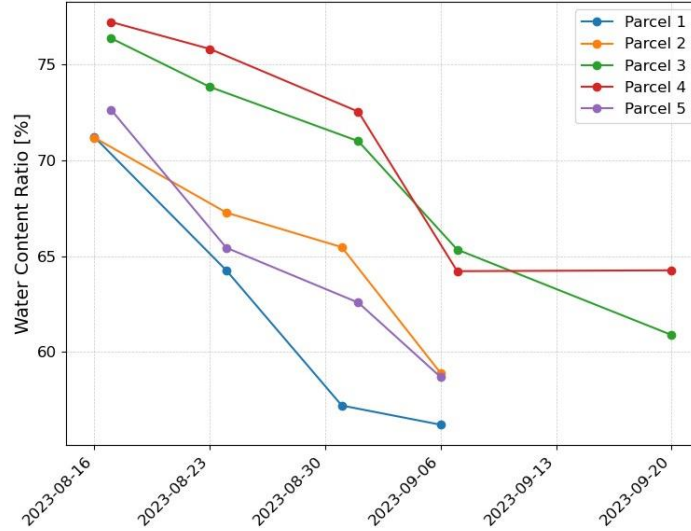


Figure 20 - Water Content Ratio [%] for the five parcels (16/08 – 09/20)

The Area Water Content per square meter has also been calculated, using the formula:

$$AWC [g/m^2] = \frac{\text{Water Content [g]}}{\text{Number of samples collected}} \times \text{Plant Density [m}^2] \quad [9]$$

The number of samples was usually 5, apart for Parcel 3 where six samples were collected. The evolution of the water content per square meter can be observed in Figure 21.

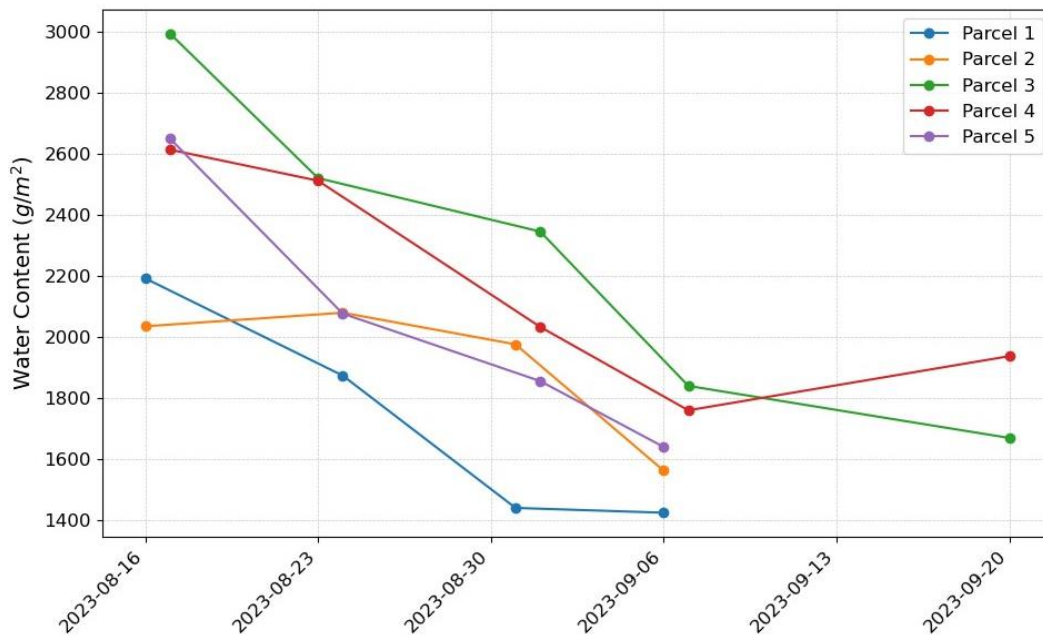


Figure 21 - Temporal evolution of the water content per square meter for the five parcels (16/08 - 20/09)

This variable proves to be particularly valuable, as it is the biophysical variable effectively influencing the backscattering signal. Consistent with expectations, there is a declining trend observed as the maize crop undergoes maturation. Parcel 3 exhibits the highest water content ratio per square meter, aligning closely with field observations.

2.4 Cobs and Plants Water Content

Cobs were separated from the stems and leaves to perform a separate analysis of moisture dynamics. The temporal evolution of their respective water content can be observed in Figure 22.

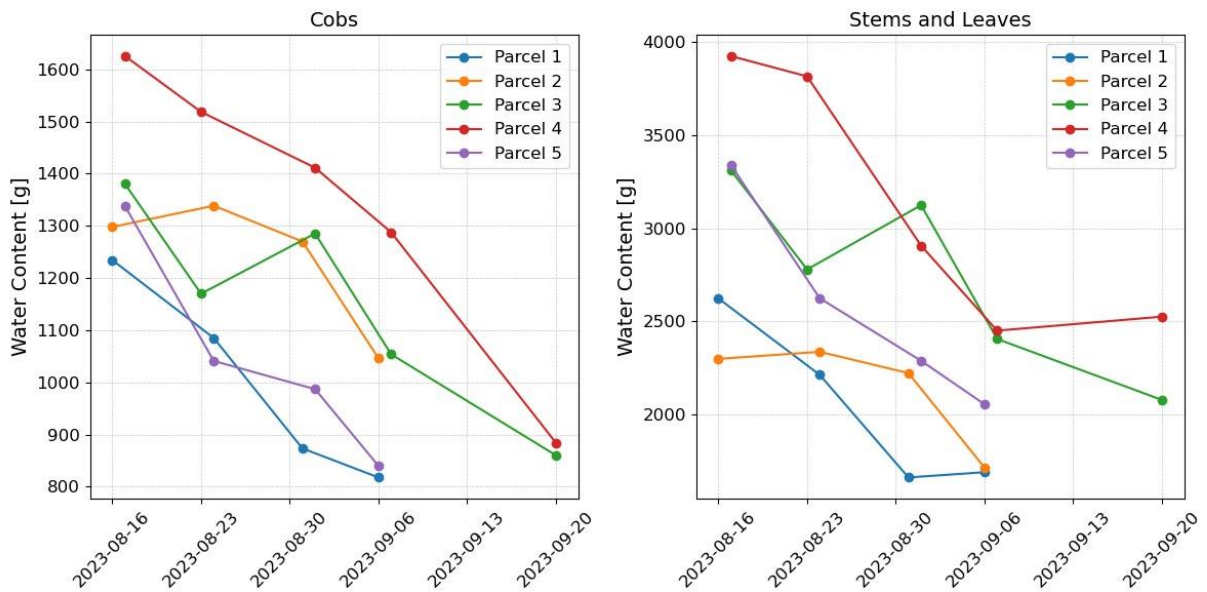


Figure 22 - Temporal evolution of the water content in cobs (left) and stems and leaves (right) of selected samples for the five parcels (16/08 – 20/09)

As expected, the water content of the cobs mirrors the water content of stems and leaves. Notably, Parcel 4 shows a distinct trend compared to the other parcels, consistent with field observations indicating that the maize plants in Parcel 4 were much greener than those in other parcels, suggesting a less advanced developmental stage.

3. SAR Data

3.1 TerraSAR-X

Figure 23 illustrates the available TerraSAR-X data points by plotting a time series of the cross-ratio in decibels. Only nine data points are available: one for Parcel 3 and Parcel 5, two for Parcel 1 and Parcel 2, and three for Parcel 4. It is worth noting that two data points for Parcel 4 were recorded on the same day, with a six-second interval, as indicated in Table 2. These two data points will be treated as a single data point for subsequent analysis, resulting in a total of eight data points available for analysis. The data points are outside the field campaign timeframe, hence do not align with the in-situ data and do not accurately represent the field conditions.

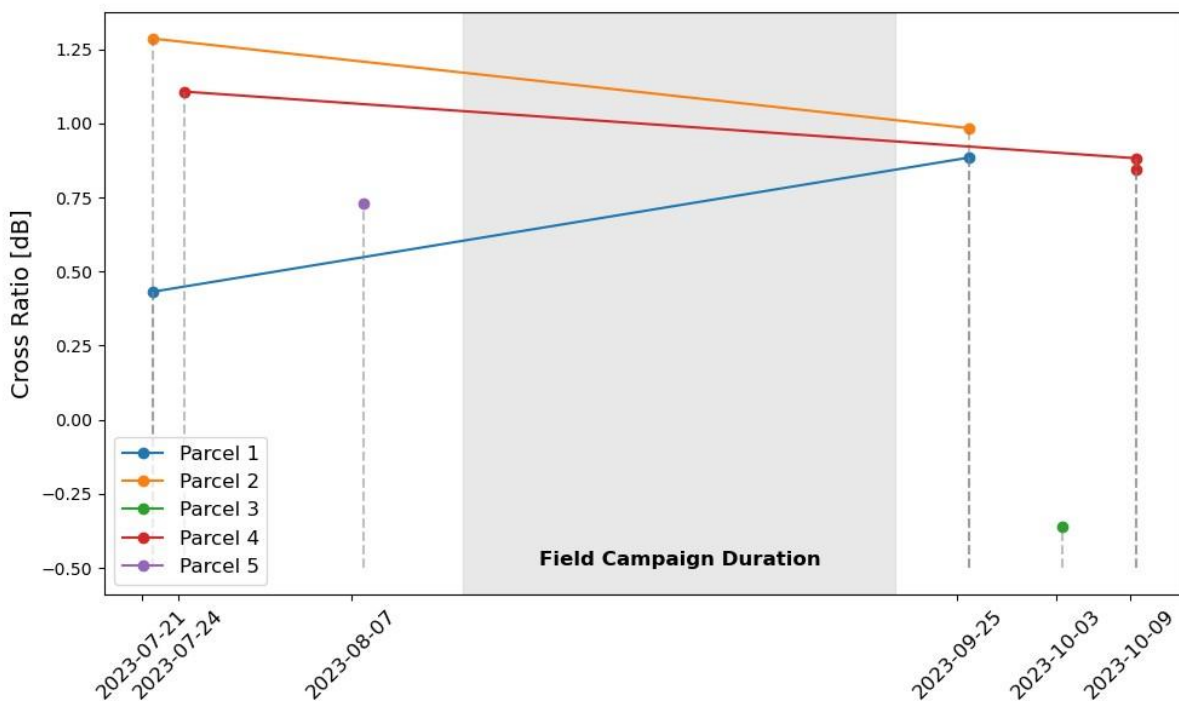


Figure 23 - TerraSAR-X Cross Ratio data points spanning the field campaign duration

3.2 Sentinel-1

Figure 24 illustrates the available Sentinel-1 data points by plotting a time series of the cross-ratio in decibels. There are significantly more data points within the duration of the field campaign, covering 4 different dates for 5 fields, totalling 20 points. As shown in Table 4, biomass was collected within a maximum of 5 days preceding or following the Sentinel-1 overpass. The Sentinel-1 data are much more aligned with the in-situ data than TerraSAR-X, as all data points are within the field campaign duration, with a maximum deviation of one day. Therefore, Sentinel-1 will be the primary satellite of reference for the final part of the results analysis.

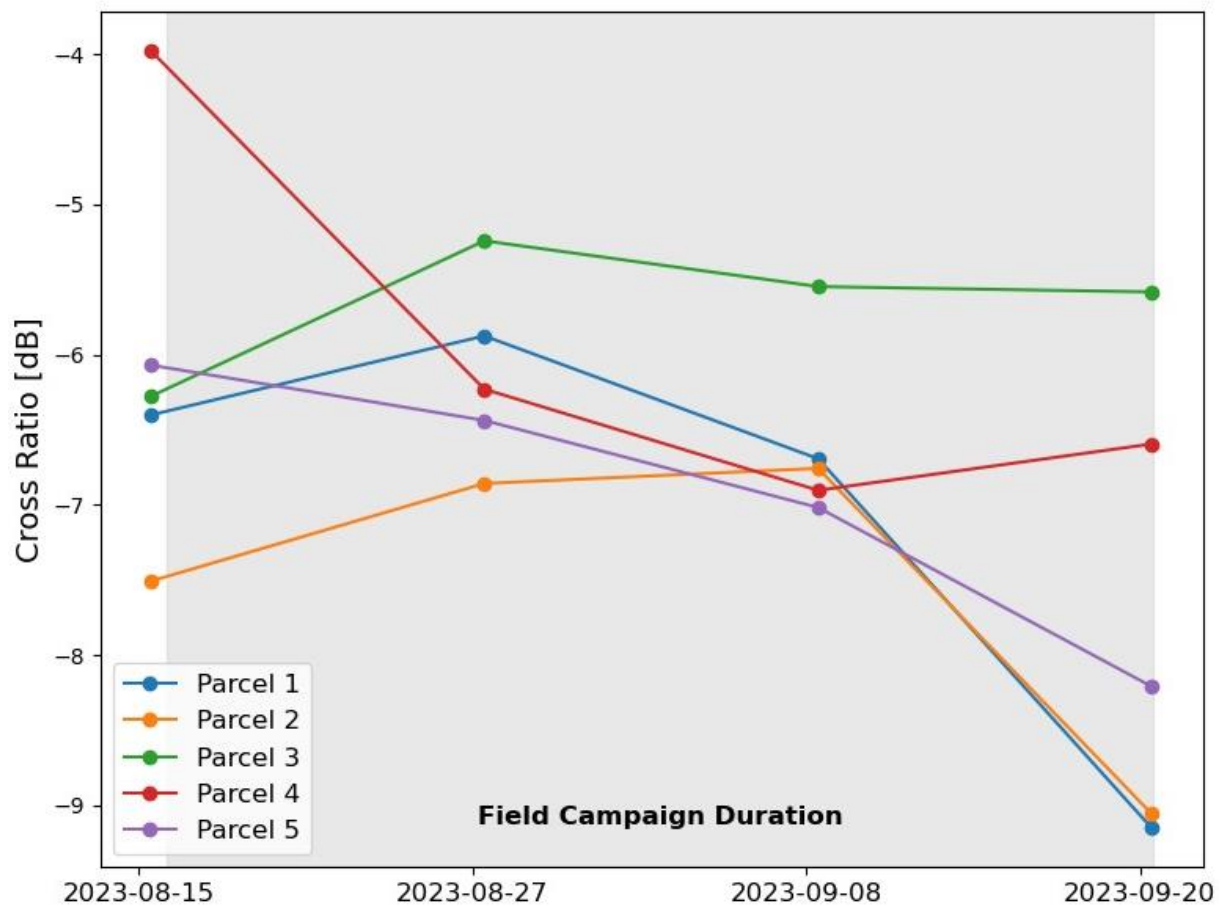


Figure 24 - Sentinel-1 Cross Ratio data points spanning the field campaign duration

4. Sensitivity Analysis

After extracting the available data points from the SAR imagery and conducting an initial exploratory analysis, the in-situ area water content data were merged with the satellite data. For each in-situ data point, the nearest concurrent satellite data point was matched. Scatterplots were then created for both TerraSAR-X and Sentinel-1 data to facilitate a comparative analysis.

4.1 TerraSAR-X

Figure 25 presents scatter plots that illustrate the relationship between TerraSAR-X data and area water content measurements for the five studied parcels. The plots display the mean backscatter values in decibels for VV (Vertical-Vertical), VH (Vertical-Horizontal)

polarisations, and the cross ratio (CR) against the area water content in grams per square meter for the five selected parcels.

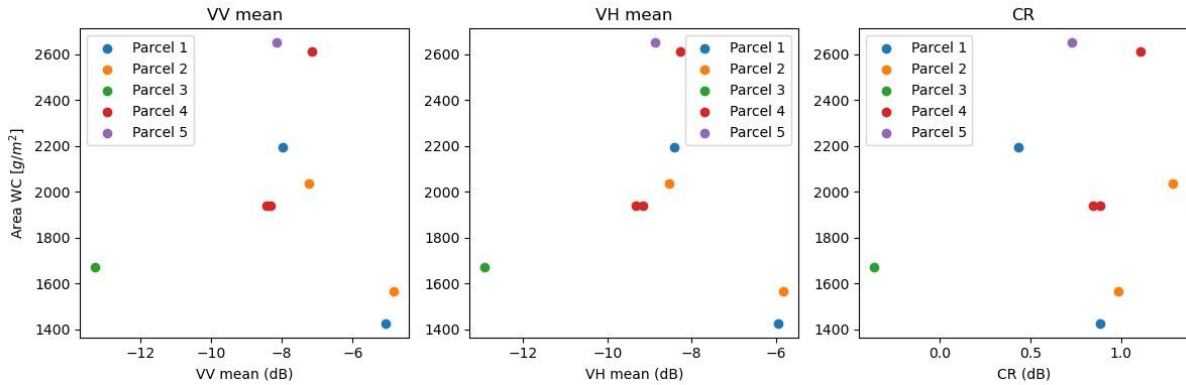


Figure 25 - Scatter plot for TerraSAR-X data merged with Area Water Content data

The scatter plots highlight how the SAR backscatter values vary with the area water content across different parcels, providing insights into the correlation between the in-situ biomass measurements and the SAR data. As highlighted in the previous section, the datapoints are very scarce and seem too few to accurately represent the field conditions. Moreover, no visible pattern or correlation is observed between the backscatter values and water content data, suggesting that it may be challenging to make effective use TerraSAR-X data in the rest of the study.

4.2 Sentinel-1

The same approach was taken for Sentinel-1 data. Figure 26 displays the mean backscatter values and the cross ratio (CR) in decibels for Sentinel-1 against the area water content in grams per square meter for the five selected parcels

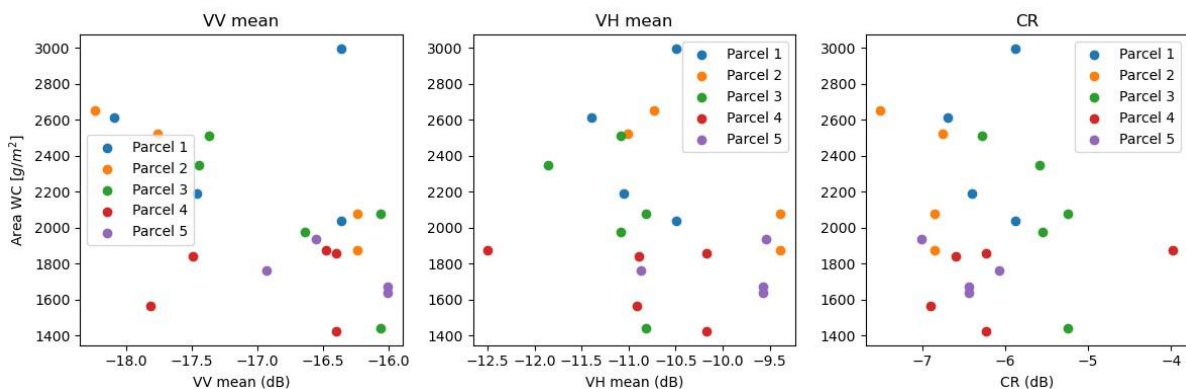


Figure 26 - Scatter plot for Sentinel-1 data merged with Area Water Content data

The number of available data points (22) allows for a more robust analysis of field conditions compared to TerraSAR-X. However, despite the increased quantity of data, clear trends or correlations between backscatter values and water content are still not apparent.

This lack of visible patterns using the available S1 data suggests that additional factors may influence the Sentinel-1 backscatter signals, or that the spatial variability within the parcels is significant. Further analysis and potentially more sophisticated statistical methods are required to uncover any underlying relationships.

5 Linear Correlation Analysis

Having established an initial exploration of the relationships between SAR data and area water content through scatter plots, linear correlation analysis is now employed. This statistical technique will quantify the strength and direction of any linear relationships that may exist between the in-situ measurements and the backscatter values obtained from the SAR imagery.

Figure 27 presents the correlation matrix between the S1 remote sensing and in-situ data, illustrating the relationships between various measured biomass parameters, including fresh biomass, dry biomass and water content in grams and percentage, area WC, cobs water content, stems and leaves WC, and the SAR backscatter values (VV mean dB, VH mean dB, cross ratio, and dpRVI). The colour gradient from blue to red represents the correlation coefficient values: strong colours representing strong correlation and lighter colours representing weak correlation and blue indicating negative correlations and red indicating positive correlations.

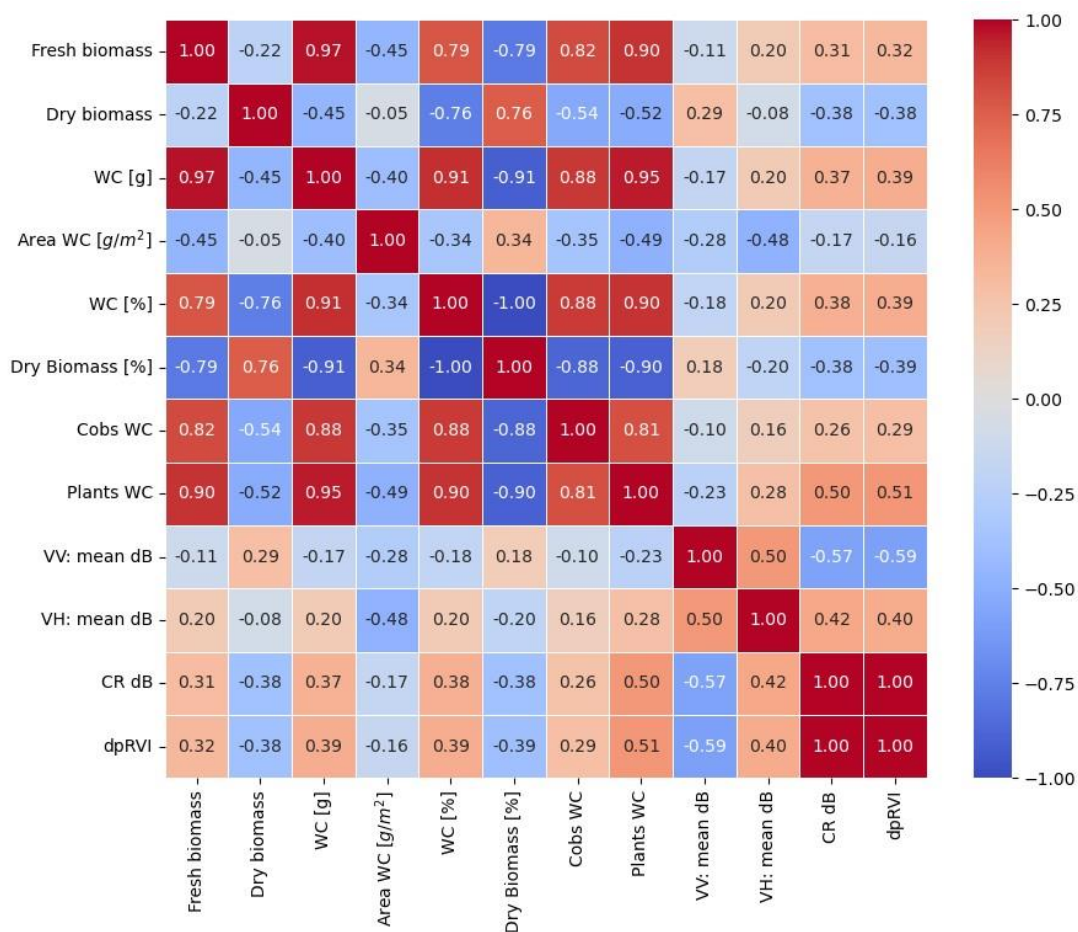


Figure 27 - Correlation matrix between in-situ- and Sentinel-1 data

As expected, strong correlations exist between the biomass parameters.

- Fresh biomass is highly positively correlated with WC (0.97), cobs WC (0.82) and plants WC (0.90). This shows that as the fresh biomass decreases, the water content in the plants decreases.
- Dry biomass is negatively correlated with WC (-0.45), cobs WC (-0.54) and plants WC (-0.52). This inverse relationship indicates that as dry biomass is increasing, the water content also decreases.
- Also, WC in grams is strongly correlated to the WC ratio (0.91), as is dry biomass in grams to its ratio (0.76).

The AWC [g/m²] shows relatively weak correlations with fresh biomass (-0.45) and total WC (-0.40). This might be surprising as one would expect a stronger relationship between the water content per area and the total water content or fresh biomass. This discrepancy might indicate that water distribution or plant density is not uniform within or across the parcels.

However, the correlations that truly matter for this study are those between the in-situ measures and the Sentinel-1 backscatter.

- The VV mean dB backscatter value displays weak negative correlations with fresh biomass (-0.11) and WC (-0.17). This suggests that the VV polarisation backscatter is only slightly influenced by water content in the field. Positive correlations are however observed with the dry biomass (0.29).
- The VH mean dB backscatter value shows weak positive correlations with fresh biomass (0.20) and WC [g] (0.20). These slightly higher values than VV indicate the higher potential to detect moisture variations in senescent maize plants. Also, the quite high negative correlation with Area WC (-0.48) might be due to the sensitivity of VH polarisation to structural changes in the canopy or variations in plant density that affect water content measurements.
- CR and dpRVI metrics both show moderate positive correlations with fresh biomass (0.31 and 0.32, respectively) and WC (0.38 and 0.39, respectively), highlighting their potential utility in assessing biomass and water content. Notably, plant WC has particularly strong correlations with CR (0.50) and dpRVI (0.51).

In order to assess the potential of TerraSAR-X data, it was also incorporated into the correlation matrix, therefore plotting an expanded correlation matrix visible at Figure 28.

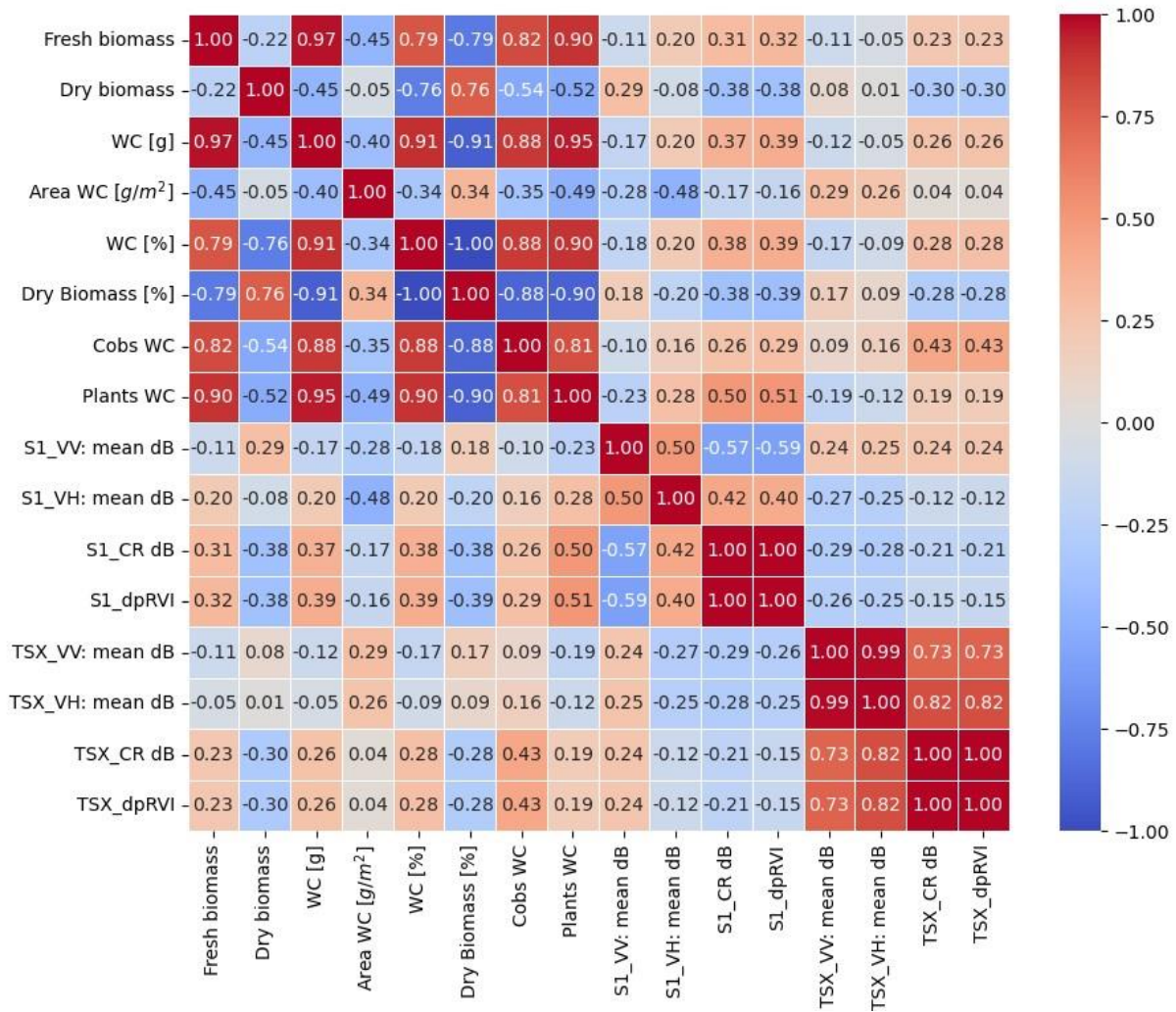


Figure 28 - Correlation matrix between in-situ, Sentinel-1 and TerraSAR-X data

While the highest correlations were observed between Cobs WC and both CR (0.43) and dpRVI (0.43), no other correlations exceeded 0.30. This suggests that, in this particular study, TerraSAR-X data provided limited utility for finding linear correlations between the in-situ and remote sensing data.

Overall, the correlation matrix from this research reveals that while some SAR backscatter values exhibit weak to moderate correlations with biomass and water content parameters, the relationships are not strong enough to draw definitive conclusions about significant correlations between the SAR signal and in-situ field conditions.

6 Non-Linear Correlation Analysis

Based on the results from the previous section, low linear correlations are observed between SAR and in-situ data. Also, the scatterplots from Vreugdenhil et al. clearly indicate the non-linear relationship between CR and WC with a large sensitivity to low WC values and saturation at high WC. This behaviour is observed for all vegetation variables and for all crop types. For these reasons, a final attempt was made to uncover potential non-linear correlations between SAR and in-situ data using a Random Forest regressor.

This non-parametric statistical method constructs an ensemble of decision trees, each modelling various aspects of the data's complexity. By averaging the predictions from multiple trees, the Random Forest regressor effectively identifies and leverages intricate, non-linear interactions between the SAR and in-situ data. This approach not only enhances predictive accuracy but also provides insights into the relative importance of different predictors, offering a robust framework for understanding the underlying patterns in the dataset (Pedregosa et al., 2011).

The dataset was organized into independent variables (predictor features) and dependent variables (target features). The predictor variables consisted of remote sensing data, specifically various Sentinel-1 backscatter coefficients (mean, standard deviation, median, minimum, and maximum), as well as CR and dpRVI. The target variables included various biomass indicators such as WC [g], WC [%], AWC [g/m²], Cobs WC [g], and Plants WC [g].

6.1 Feature Importance Scores

The Random Forest Regressor is applied using all the default parameters provided by Scikit-learn. The aim of the model was to determine if any combination of remote sensing data could potentially reveal certain correlations with biomass parameters. The model was fit to the data to learn the underlying patterns and relationships between the predictor variables and the target variables. For each Random Forest, feature importance scores were calculated, indicating the contribution of each predictor variable to the model's predictions. These scores were used to identify the most influential variables and understand the underlying non-linear relationships.

Initially, Random Forests were performed using all available Sentinel-1 backscatter coefficients as predictor variables. However, after running the Random Forests multiple times, it was noticed that the feature importance scores were highest for the minimum and maximum backscatter coefficients, as shown in Table 14.

Table 14 - Feature Importance Score for all Sentinel-1 backscatter coefficients and indexes on WC [g] and AWC [g/m²]

Feature	Importance score for WC [g]	Importance Score for AWC [g/m ²]
VV: mean	0.0422	0.0145
VV: std	0.0559	0.0125
VV: median	0.0377	0.1305
VV: min	0.2500	0.3218
VV: max	0.1138	0.0263
VV: mean dB	0.0430	0.0135
VH: mean	0.0305	0.0762
VH: std	0.0635	0.1462
VH: median	0.0776	0.0426
VH: min	0.1003	0.0267
VH: max	0.0560	0.1050
VH: mean dB	0.0277	0.0481
CR	0.0507	0.0152
dpRVI	0.0511	0.0210

It became evident that the minimum and maximum backscatter coefficients had the highest feature importance scores for both WC [g] and AWC [g/m²]. This suggests that these particular coefficients significantly influenced the model's predictions.

However, using extreme values such as minimum and maximum backscatter coefficients can sometimes lead to overfitting and unreliable results due to the potential presence of noise and outliers in the analysed data. To mitigate this issue and obtain more robust and generalisable results, it was decided to only focus on the mean and standard deviation backscatter coefficients, as well as the mean in decibels, CR, and dpRVI, to evaluate the correlation between SAR and in-situ variables.

6.2 Performance metrics

Several evaluation metrics were calculated to evaluate the model's performance: R-squared (R²), Root Mean Squared Error (RMSE), Relative Root Mean Squared Error (RRMSE), and the Out-of-Bag (OOB) score.

- **R-squared (R²):** This metric measures the proportion of the variance in the target variable that is predictable from the predictor variables. Higher R² values indicate a better fit of the model to the data.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad [10]$$

With:

- y_i = observed values ;
- $\hat{y}_i$ = predicted values ;
- $\bar{y}$ = mean of the observed values ;

- n = number of observations

- **Root Mean Squared Error (RMSE):** This metric provides a measure of the average magnitude of the prediction errors, giving a sense of how close the predicted values are to the observed values. Lower RMSE values indicate better model accuracy.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad [11]$$

- **Relative Root Mean Squared Error (RRMSE):** This is the RMSE expressed as a percentage of the mean of the observed data, providing a normalized measure of prediction error. Lower RRMSE values indicate better model performance relative to the scale of the data.

$$RRMSE = \frac{RMSE}{\bar{y}} \times 100 \quad [12]$$

- **Out-of-Bag (OOB) Score:** This is an internal validation score used in Random Forest models. It is computed using the predictions for each sample from the trees that did not include that sample in their bootstrap sample. Higher OOB scores indicate better model performance and generalisation to unseen data. It is therefore a useful indicator to evaluate how well the model is going to perform with new data.

These performance metrics collectively provide a comprehensive assessment of the model's performance and its potential to capture non-linear relationships between SAR data and in-situ measurements.

6.3 GridSearchCV and Hyperparameters

To optimise the Random Forest model, *GridSearchCV* from Scikit-learn was employed. GridSearchCV is a technique that performs an exhaustive search over a specified parameter grid to find the best combination of hyperparameters for a model. This method involves training the model with different sets of hyperparameters and evaluating its performance using cross-validation to ensure robust results (Great Learning Team, 2023).

GridSearchCV was used to define a function to minimise the absolute difference between R^2 and the OOB score. By focusing on minimizing the difference between these two metrics, the goal was to balance the model's fit and its ability to generalise to new data.

The following hyperparameters were tuned during the GridSearchCV process:

- *max_depth* : This parameter specifies the maximum depth of the individual trees in the Random Forest. Limiting the depth of the trees helps prevent overfitting by restricting the complexity of the model. A deeper tree might capture more details of the training data but can also capture noise, leading to poor generalization. Max depth was chosen between 10, 20, or 30.
- *max_features* : This parameter determines the number of features to consider when looking for the best split at each node. Setting *max_features* to 1.0 means that all features are

considered, while 'sqrt' and 'log2' are other commonly used settings that limit the number of features, potentially reducing overfitting and improving generalization.

- *n_estimators* : This parameter specifies the number of trees in the forest. A higher number of trees can lead to more stable and accurate predictions, as the model averages the results from multiple trees. However, increasing the number of trees also increases computational complexity. Important to note that some inputs did not have oob scores when too few trees were used for determination. This is why *n_estimators* were chosen between 50, 75, 100 or 200 trees.

6.4 Leave-one-out Cross-Validation

Finally, cross-validation was performed to assess the generalisation capability of the model and ideally prevent overfitting by evaluating its performance on an independent dataset. In this case, Leave-One-Out Cross-Validation (LOOCV) was used, a specific type of cross-validation where the model is trained on all data points except one, which is used as a test set. This process is repeated for each data point, ensuring that each point is used as a test set once. LOOCV is particularly useful for small datasets as it maximises the use of available data without being too computationally intensive. For this cross-validation method, the Root Mean Squared Error (RMSE) and Relative Root Mean Squared Error (RRMSE) were calculated to analyse its generalisation capability, offering insights into the model's accuracy and its ability to generalise to new, unseen data (scikit-learn.org, s. d.).

6.5 Results for WC [g] and AWC [g/m²]

The following sections present the results for WC [g] and AWC [g/m²], detailing the hyperparameters used to optimise the model, feature importance scores, and performance metrics. The results for the LOOCV (Leave-One-Out Cross-Validation) are also included.

6.5.1 WC [g]

Table 15 - Feature Importance Scores for WC [g]

Feature	Importance score
VV: mean	0.1053
VV: std	0.1664
VV: mean dB	0.1363
VH: mean	0.0761
VH: std	0.1972
VH: mean dB	0.0569
CR	0.1287
dpRVI	0.1331

From Table 15, it appears that the VH: std feature has the highest importance score (0.1972), followed closely by VV: std (0.1664). This indicates that the variability in the backscatter values (as measured by the standard deviation) might be a significant predictor for water content. The backscatter values for the VV polarization generally have slightly higher feature importance scores compared to the VH polarization. However, no single feature overwhelmingly dominates.

The selected hyperparameters determined by GridSearchCV were $\text{max_depth} = 10$, $\text{max_features} = 1.0$, and $\text{n_estimators} = 100$. These settings were chosen to minimise the difference between R^2 and oob.

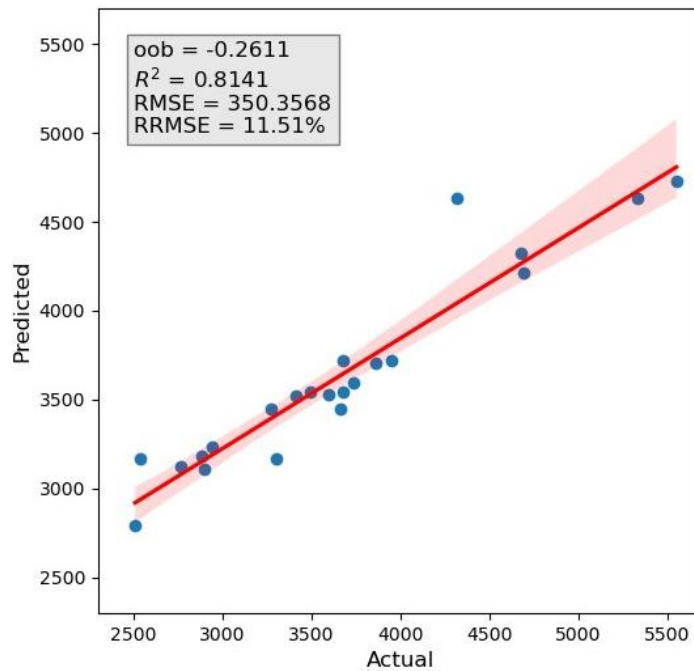


Figure 29 - Actual vs predicted WC [g] using the optimised Random Forest model

Figure 29 illustrates the relationship between the actual and predicted WC [g] values using the optimized Random Forest model. The plot shows a clear positive correlation, indicating that the model predictions align well with the observed values. The red line represents the line of best fit, with the shaded area showing the confidence interval.

These metrics indicate that the model has a good fit, with an R^2 value of 0.8141, meaning that approximately 81% of the variance in WC [g] is explained by the model. The RMSE value of 350.3568 suggests that the average prediction error is around 350 grams, while the RRMSE of 11.51% indicates that this error is relatively low compared to the mean of the observed data.

Cross-validation was then performed on the fitted model to assess the generalization capability of the model and ideally prevent overfitting by evaluating its performance on an independent dataset. The scatter plot presented at Figure 30 compares the actual water content (WC [g]) values against the predicted values obtained from the LOOCV method. The red dashed line

represents the line of perfect prediction ($y=x$), where predicted values exactly match the actual values.

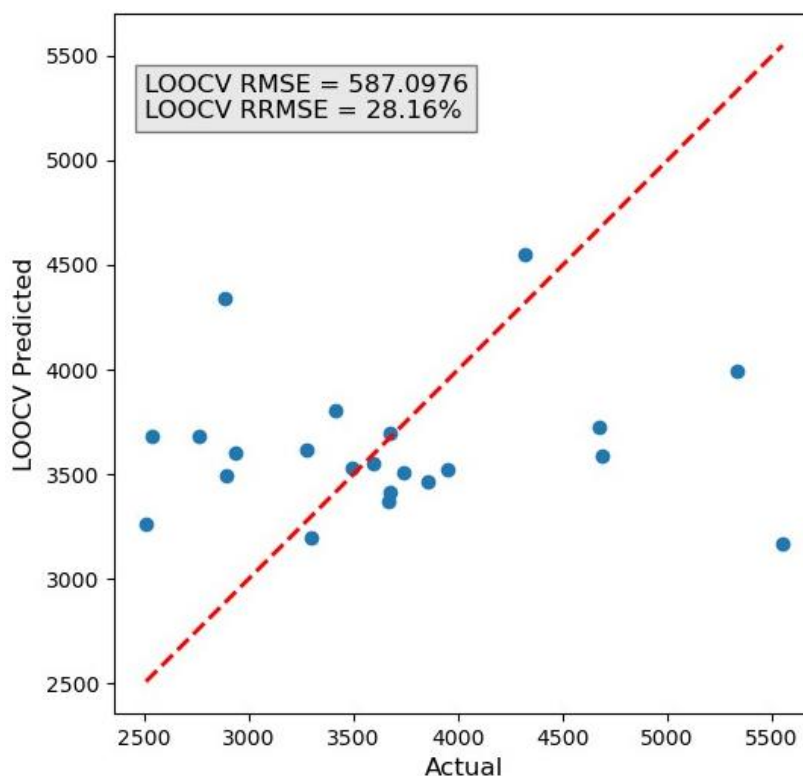


Figure 30 - LOOCV Results for WC [g]

The model's LOOCV performance is quantified by the Root Mean Squared Error (RMSE) of 587.09976 and Relative Root Mean Squared Error (RRMSE) of 28.16%, indicating that the model has limited generalisation capability and performs poorly on unseen data.

In conclusion, the model fits very well with the training data, but performs poorly on unseen water content data. Therefore, the model is overfitted and has no reliable prediction capability.

6.5.2 AWC [g/m²]

Table 16 - Feature Importance Scores for AWC [g/m²]

Feature	Importance score
VV: mean	0.0725
VV: std	0.0873
VV: mean dB	0.1059
VH: mean	0.1187
VH: std	0.3851
VH: mean dB	0.1320
CR	0.0346
dpRVI	0.0640

Table 16 shows that the VH polarisation has notably higher scores than the VV polarisation. VH: std has the highest importance score (0.03851), followed by VH: mean dB (0.1320). The higher importance of these VH features could potentially indicate that VH polarisation is slightly better at detecting area water content compared to VV polarisations.

The hyperparameters optimized by GridSearchCV were max_depth = 10, max_features = 1.0, and n_estimators = 50, aimed at minimizing the discrepancy between R^2 and the oob score.

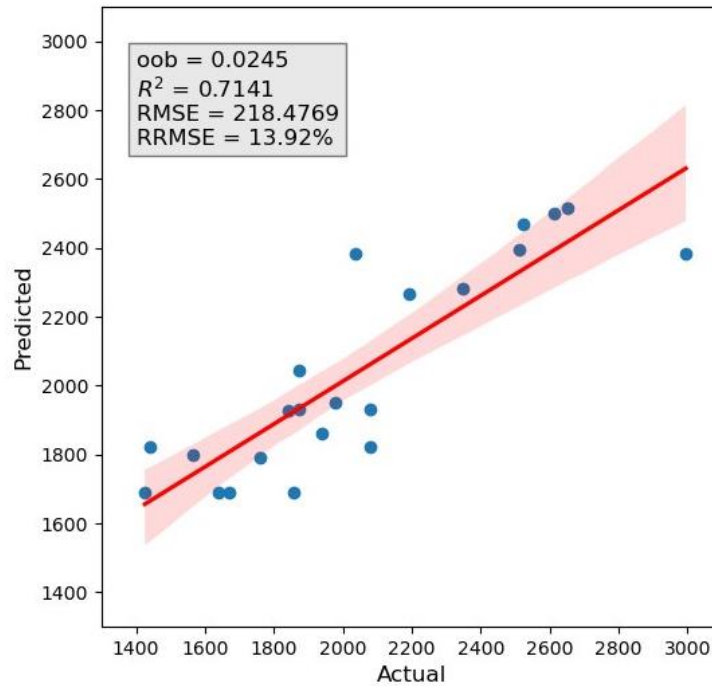


Figure 31 - Actual vs Predicted AWC [g/m²] using the optimised Random Forest model

Figure 31 illustrates a positive correlation between actual and predicted AWC [g/m²] values using the optimised Random Forest model. The R^2 value of 0.7141 indicates that about 71% of the variance in AWC [g/m²] is captured by the model. An RMSE of 218.4769 suggests that the average prediction error is around 218 grams per square meter, and an RRMSE of 13.92% shows a moderate error relative to the observed data mean.

In the same manner as for WC [g], LOOCV was also performed. Figure 32 compares actual AWC [g/m²] values with the LOOCV predictions.

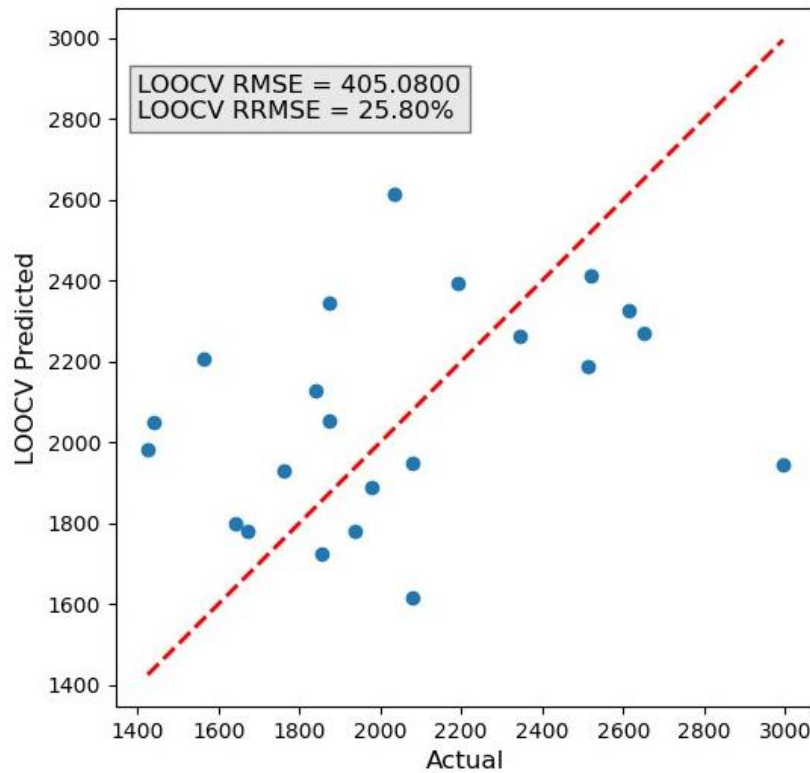


Figure 32 - LOOCV results for AWC [g/m²]

The LOOCV performance, with an RMSE of 405.0800 and RRMSE of 25.80%, indicates limited generalisation capability and suggests the model performs poorly on unseen data.

Comparing the results for WC [g] and AWC [g/m²], it is notable that the LOOCV RRMSE is lower for AWC (25.80%) compared to WC (28.16%). This could suggest that the model generalises slightly better for AWC than for WC, although points seem to be situated further from the line of perfect predictions. One possible reason for this difference is that AWC, being a measure of area water content, might have a more stable and predictable relationship with the Sentinel-1 backscatter features. Additionally, the lower number of trees (50) used for AWC could have helped in preventing overfitting to some extent, compared to the higher number of trees (100) used for WC.

In conclusion, while the model fits reasonably well with the training data for AWC [g/m²], it shows poor performance on new, unseen data, indicating clear overfitting and a lack of robust prediction capability. Similar to the results for WC [g], the generalisation capability of the model remains a challenge, highlighting the need for further refinement and possibly the incorporation of additional data with higher correlations.

Chapter 6: Discussion

As expected, water content in maize plants decreased over time during the field campaign. The average area water content for the five parcels was 2497.99 grams per square meter on day 1 (16 August) and 1646.17 grams per square meter at the end of the campaign (20 September). Although three of the five parcels had already been harvested by the final field visit, the remaining fields showed a similar trend of desiccation. Gao et al. (2023) observed that the irreversible turning point leading to organ death in maize is associated with a minimum threshold of about 60% plant water content, which corresponds to the optimal time for maize harvesting. The average water content ratio at which maize was harvested by the farmers during the field campaign was 57.91%, aligning with this observation. Therefore, we can conclude that the biomass data collected during the field campaign are representative of the typical desiccation process in senescent maize plants.

After an exploratory analysis of the available SAR data and plotting the various data points for TerraSAR-X and Sentinel-1, it appeared that only eight data points were available for TSX, compared to 20 for S1. Additionally, all TSX data points were outside the field campaign duration, with the closest date being 10 days before the campaign and the farthest 32 days after. Therefore, it was clear that TSX would not accurately represent the field conditions due to its lack of alignment with the in-situ measurements, making it less valuable for this research. Consequently, the focus shifted to S1 data.

Vreugdenhil et al. (2018) plotted a scatterplot of S1 backscatter with in-situ measured VWC [kg/m²] (which corresponds in this work to the AWC [g/m²]) for three different crops over their entire growing stage: wintercereal, rape and corn (Figure 33). Emerging trends were observed for CR and VH polarisations. In contrast, the scatter plots produced for this research contained no visible trends, suggesting that S1 may not be able to detect variations in plant water content effectively.

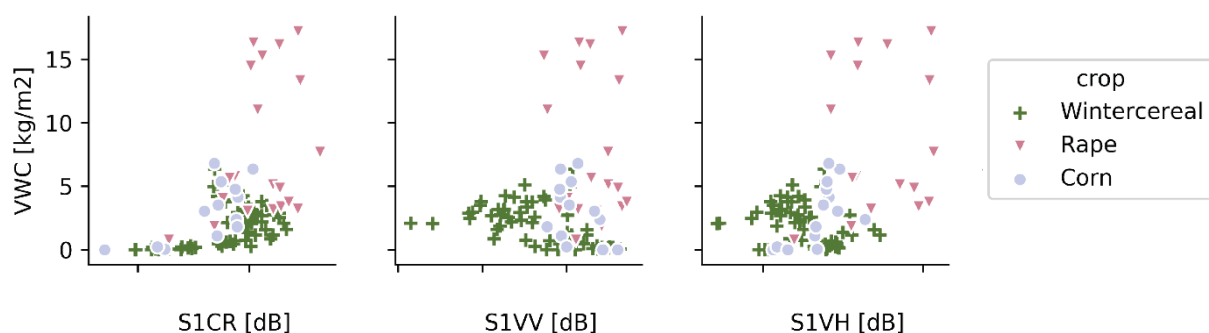


Figure 33 - Scatter plot of S1 CR (left), VV (middle) and VH (right) backscatter with in-situ measured VWC [kg/m²]. Corn measures are indicated in blue. (Vreugdenhil et al. (2018))

A linear correlation analysis using a correlation matrix between the various in-situ- and satellite data provided limited insights. Correlations with TSX data were significantly low (<0.30), leading to an analysis of S1 correlations exclusively. As expected, strong correlations were observed between the various biomass and water content parameters. However, the correlations that truly matter – between the in-situ measurements and the radar backscatter values – showed weak to moderate correlations with biomass and content parameters. The highest correlation was found between AWC [g/m²] and VH: mean dB (-0.47), and between Plants WC and CR and dpRVI (0.50). Other correlations remained relatively low (<0.40). Those results can be compared to linear model results from Vreugdenhil et al. (2018). In their study, they separated maize stems from leaves. Based on their linear model, CR showed the highest correlation to AWC with $R^2 = 0.48$, followed by VH (0.44) and VV (0.15). They also observed that VH backscatter is most sensitive to WC in leaves ($R^2 = 0.49$). Conversely, VV backscatter appears to be more sensitive to soil backscatter ($R^2 = 0.24$).

Finally, a non-linear correlation analysis was performed on WC [g] and AWC [g/m²] using mean and std backscatter coefficients, as well as CR and dpRVI for S1 data. Feature importance scores were computed using only CR, dpRVI and the mean and standard deviation backscatter coefficients. It was found that VH backscatter coefficients had higher importance scores than for the VV polarisation, especially for AWC (VH:std = 0.3851 , VH:mean dB = 0.1320). This aligns with results from other studies (Vreugdenhil et al., 2018; Khabbazan et al., 2019) stating that VH polarisation is better in capturing water content variations in crops. However, the feature importance score for CR ($=0.0346$) was particularly low compared to the results from Vreugdenhil et al., where the feature importance of CR equals 0.35 . Additionally, after running the model through leave-one-out cross-validation, it was found that although the model fits quite well to the training data, its generalisation capability to new data was virtually non-existent.

Therefore, the approach pursued in this work proves ineffective for accurate plant water content monitoring in senescent maize, in this particular area. Several factors might contribute to these limitations:

One significant factor is the potential flaws in the in-situ data collection. The limited number of samples and the possibility of measurement errors could affect the reliability of the correlation between in-situ measurements and satellite data. Inaccurate or insufficient in-situ data can lead to misleading conclusions about the effectiveness of the remote sensing techniques used.

Another contributing factor is the small size of the parcels used in this study, which may limit the representativity of radar measurements. Smaller areas are more susceptible to border effects and variability within the parcel, distorting the correlation between radar backscatter and plant

water content. This lack of representativity can undermine the applicability of the findings to larger agricultural areas, reducing the overall effectiveness of the study.

Additionally, there is a possibility of radar signal saturation, especially at higher biomass levels such as in maize fields. When the vegetation is dense, the radar signal may not penetrate deeply enough to provide accurate information about the water content within the plants. This saturation effect can result in weaker correlations between the radar backscatter values and the actual water content of the maize plants, limiting the ability of the radar to accurately monitor variations in water content.

Chapter 7: Conclusion

This study aimed to evaluate the feasibility of a plant water content retrieval method in senescent maize crops using multi-frequency (C- and X-band) SAR backscatter data. To achieve this, an intensive field campaign was conducted in the Taiyuan Basin, Northern China, with the help of our Chinese partners from the Taiyuan University of Technology. These efforts were intended to gather comprehensive in-situ data on the final developmental stages of maize crops.

During the field campaign, biophysiological data were collected in five parcels through five visits from mid-August to late September 2023. Temporal evolution of variables such as fresh biomass, dry biomass, and water content were documented, providing insights into the drying process of maize during senescence. Subsequent analysis included correlating this in-situ data with SAR data from TerraSAR-X and Sentinel-1. The TerraSAR-X dataset proved inadequate due to sparse data points misaligned with the field campaign. Despite better temporal alignment, Sentinel-1 data also failed to show clear trends or significant correlations with in-situ measurements. Linear and non-linear correlation analyses were conducted to explore relationships between SAR backscatter values and the collected biophysiological parameters. Some trends similar to those found in the literature were observed; however, these findings must be interpreted with caution as the exact representativity of the satellite data in relation to the in-situ measurements remains uncertain. Consequently, the current method proved ineffective for accurate plant water content monitoring in senescent maize.

Several improvements could enhance the method's effectiveness:

First, improving field campaign resources might be necessary. A larger dataset encompassing a wider range of field conditions could potentially lead to more robust correlations. Increasing the number of sampled parcels and the frequency of measurements would provide a more comprehensive dataset for analysis, capturing more variability and thus offering better insights into the relationship between SAR backscatter and maize water content.

Second, selecting larger parcels with homogeneous field practices could address the issue of spatial variability within small parcels. Often, these small parcels are composed of multiple fields managed by different owners with varying cultivation methods, which may limit the correlation analysis. By choosing larger parcels with uniform cultivation practices, the variability within each parcel can be reduced, thereby improving the reliability and consistency of the results.

Third, selecting areas with delayed harvest dates could provide more data points during the critical period of maize senescence. This would allow for the collection of data over a longer period, providing a more detailed time series of water content variations.

Finally, using other radar bands, such as the L-band, could enhance the detection of water content in maize. The X-band, with its sensitivity to surface characteristics, may be less

effective for monitoring volumetric changes in senescent maize. The C-band, which strikes a balance between penetration and surface sensitivity, has shown some promise, but its effectiveness can still be limited in dense biomass conditions. In contrast, the L-band, which is less likely to be saturated, can penetrate deeper into the canopy and offer better insights into plant water content. This deeper penetration could lead to more accurate assessments of water content, especially in dense biomass crops such as maize.

A more comprehensive approach to field campaign resources, parcel selection, and remote sensing bands could therefore significantly improve the monitoring of plant water content in senescent maize. In a broader context, developing reliable methods for monitoring the optimal time of harvest could serve as a valuable decision support tool for crop harvesting order. Accurate predictions of plant water content can enhance harvest logistics, allowing for better resource management and ultimately contributing to increased agricultural efficiency and productivity. While this study faced challenges, it underlines the potential benefits of integrating remote sensing technology with field measurements to advance precision agriculture.

References

- Abendroth, L., Elmore, R., Hartzler, R. G., McGrath, C., Mueller, D. S., Munkvold, G. P., Pope, R., Rice, M. E., Robertson, A. E., & Sawyer, J. (2009). Corn field guide. *Iowa State University, Extension Service: Ames, IA, USA*. <https://core.ac.uk/download/pdf/38931312.pdf>
- Afuakwa, J. J., & Crookston, R. K. (1984). Using the Kernel Milk Line to Visually Monitor Grain Maturity in Maize1. *Crop Science*, 24(4), cropscl1984.0011183X002400040015x. <https://doi.org/10.2135/cropscl1984.0011183X002400040015x>
- Bal, M. A., Coors, J. G., & Shaver, R. D. (1997). Impact of the Maturity of Corn for Use as Silage in the Diets of Dairy Cows on Intake, Digestion, and Milk Production. *Journal of Dairy Science*, 80(10), 2497-2503. [https://doi.org/10.3168/jds.S0022-0302\(97\)76202-7](https://doi.org/10.3168/jds.S0022-0302(97)76202-7)
- Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., Weiss, M., Samain, O., Roujean, J. L., & Leroy, M. (2007). LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION : Part 1: Principles of the algorithm. *Remote Sensing of Environment*, 110(3), 275-286. <https://doi.org/10.1016/j.rse.2007.02.018>
- Bargiel, D. (2017). A new method for crop classification combining time series of radar images and crop phenology information. *Remote Sensing of Environment*, 198, 369-383. <https://doi.org/10.1016/j.rse.2017.06.022>
- Bazzi, H., Baghdadi, N., El Hajj, M., Zribi, M., Minh, D. H. T., Ndikumana, E., Courault, D., & Belhoucette, H. (2019). Mapping Paddy Rice Using Sentinel-1 SAR Time Series in Camargue, France. *Remote Sensing*, 11(7), Article 7. <https://doi.org/10.3390/rs11070887>
- Blumberg, D. (2007). High resolution X-band SAR imagery for precise agriculture and crop monitoring. *European Space Agency, (Special Publication) ESA SP*.
- Bontemps, S., Arias, M., Cara, C., Dedieu, G., Guzzonato, E., Hagolle, O., Inglada, J., Matton, N., Morin, D., Popescu, R., Rabaute, T., Savinaud, M., Sepulcre, G., Valero, S., Ahmad, I., Bégué, A., Wu, B., De Abelleira, D., Diarra, A., ... Defourny, P. (2015). Building a Data Set over 12 Globally Distributed Sites to Support the Development of Agriculture Monitoring Applications with Sentinel-2. *Remote Sensing*, 7(12), Article 12. <https://doi.org/10.3390/rs71215815>
- Bouchat, J., Tronquo, E., Orban, A., Neyt, X., Verhoest, N. E. C., & Defourny, P. (2022). Green Area Index and Soil Moisture Retrieval in Maize Fields Using Multi-Polarized C- and L-Band SAR Data and the Water Cloud Model. *Remote Sensing*, 14(10), Article 10. <https://doi.org/10.3390/rs14102496>
- Boxer, B. (2024, mars 20). *Shanxi | Province, Location, Climate, Population, & Facts | Britannica*. <https://www.britannica.com/place/Shanxi>

Brown, S. C. M., Cookmartin, G., Morrison, K., McDonald, A. J., Quegan, S., Anderson, C., Cordey, R., & Dampney, P. (2000). *Wheat scattering mechanisms observed in near-field radar imagery compared with results from a radiative transfer model*. 7, 2933-2935. Scopus.

Brown, S. C. M., Quegan, S., Morrison, K., Bennett, J. C., & Cookmartin, G. (2003). High-resolution measurements of scattering in wheat canopies—Implications for crop parameter retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, 41(7 PART I), 1602-1610. Scopus. <https://doi.org/10.1109/TGRS.2003.814132>

Bush, T. F., & Ulaby, F. T. (1978). An evaluation of radar as a crop classifier. *Remote Sensing of Environment*, 7(1), 15-36. [https://doi.org/10.1016/0034-4257\(78\)90004-4](https://doi.org/10.1016/0034-4257(78)90004-4)

Capodici, F., D'Urso, G., & Maltese, A. (2013). Investigating the Relationship between X-Band SAR Data from COSMO-SkyMed Satellite and NDVI for LAI Detection. *Remote Sensing*, 5(3), Article 3. <https://doi.org/10.3390/rs5031389>

Doraiswamy, P. C., Hatfield, J. L., Jackson, T. J., Akhmedov, B., Prueger, J., & Stern, A. (2004). Crop condition and yield simulations using Landsat and MODIS. *Remote Sensing of Environment*, 92(4), 548-559. <https://doi.org/10.1016/j.rse.2004.05.017>

Earth Science Data Systems, N. (2020, avril 10). *What is Synthetic Aperture Radar? | Earthdata [Backgrounder]*. Earth Science Data Systems, NASA. <https://www.earthdata.nasa.gov/learn/backgrounders/what-is-sar>

Echoes in space – EO College. (s. d.). Consulté 5 novembre 2023, à l'adresse <https://eo-college.org/courses/echoes-in-space/>

El Hajj, M., Baghdadi, N., Bazzi, H., & Zribi, M. (2019). Penetration Analysis of SAR Signals in the C and L Bands for Wheat, Maize, and Grasslands. *Remote Sensing*, 11(1), Article 1. <https://doi.org/10.3390/rs11010031>

Erenstein, O., Chamberlin, J., & Sonder, K. (2021). Estimating the global number and distribution of maize and wheat farms. *Global Food Security*, 30, 100558. <https://doi.org/10.1016/j.gfs.2021.100558>

Erenstein, O., Jaleta, M., Sonder, K., Mottaleb, K., & Prasanna, B. M. (2022). Global maize production, consumption and trade: Trends and R&D implications. *Food Security*, 14(5), 1295-1319. <https://doi.org/10.1007/s12571-022-01288-7>

Filya, I. (2004). Nutritive value and aerobic stability of whole crop maize silage harvested at four stages of maturity. *Animal Feed Science and Technology*, 116(1), 141-150. <https://doi.org/10.1016/j.anifeedsci.2004.06.003>

Gao, S., Ming, B., Li, L.-L., Wang, Y.-Z., Xue, J., Hou, P., Wang, K.-R., Zhou, S.-L., Li, S.-K., & Xie, R.-Z. (2023). Temporal and Spatial Dynamics of Organ Water Content in Maize with Different Senescence Types. *Plants*, 12(18), 3269. <https://doi.org/10.3390/plants12183269>

Great Learning Team. (2023, mai 30). *Hyperparameter Tuning with GridSearchCV*. Great Learning Blog: Free Resources What Matters to Shape Your Career! <https://www.mygreatlearning.com/blog/gridsearchcv/>

Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., & Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies : Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, 90(3), 337-352. <https://doi.org/10.1016/j.rse.2003.12.013>

Hanway, J. J. (1963). Growth Stages of Corn (*Zea mays*, L.)1. *Agronomy Journal*, 55(5), 487-492. <https://doi.org/10.2134/agronj1963.00021962005500050024x>

Hatew, B., Bannink, A., Laar, H. van, Jonge, L. H. de, & Dijkstra, J. (2016). Increasing harvest maturity of whole-plant corn silage reduces methane emission of lactating dairy cows. *Journal of Dairy Science*, 99(1), 354-368. <https://doi.org/10.3168/jds.2015-10047>

Heiniger, R. (2018, août 17). *Harvesting Corn : What Grain Moisture Should I Harvest Corn At?* <https://corn.ces.ncsu.edu/2018/08/harvesting-corn-what-grain-moisture-should-i-harvest-corn-at/>

Hosseini, M., & McNairn, H. (2017). Using multi-polarization C- and L-band synthetic aperture radar to estimate biomass and soil moisture of wheat fields. *International Journal of Applied Earth Observation and Geoinformation*, 58, 50-64. <https://doi.org/10.1016/j.jag.2017.01.006>

Hosseini, M., McNairn, H., Robertson, L. D., Xianfeng, J., Merzouki, A., & Mitchell, S. (2018). *Soil Moisture and Vegetation Measurement Guideline*.

Introduction to SAR - HyP3. (s. d.). Consulté 7 juin 2024, à l'adresse https://hyp3-docs.asf.alaska.edu/guides/introduction_to_sar/

Jensen, C., Weisbjerg, M. R., Nørgaard, P., & Hvelplund, T. (2005). Effect of maize silage maturity on site of starch and NDF digestion in lactating dairy cows. *Animal Feed Science and Technology*, 118(3), 279-294. <https://doi.org/10.1016/j.anifeedsci.2004.10.011>

Johnson, L., Harrison, J. H., Hunt, C., Shinnars, K., Doggett, C. G., & Sapienza, D. (1999). Nutritive Value of Corn Silage as Affected by Maturity and Mechanical Processing: A Contemporary Review. *Journal of Dairy Science*, 82(12), 2813-2825. [https://doi.org/10.3168/jds.S0022-0302\(99\)75540-2](https://doi.org/10.3168/jds.S0022-0302(99)75540-2)

Jongschaap, R., & Schouten, L. (2005). Predicting wheat production at regional scale by integration of remote sensing data with a simulation model. *Agronomy for Sustainable Development*, 25(4), 481-489.

Khabbazan, S., Vermunt, P., Steele-Dunne, S., Arntz, L., Marinetti, C., Valk, D., Iannini, L., Molijn, R., Westerdijk, C. E., & Sande, C. (2019). Crop Monitoring Using Sentinel-1 Data : A Case Study from The Netherlands. *Remote Sensing*, 11, 1887. <https://doi.org/10.3390/rs11161887>

Kim, Y., & Van Zyl, J. J. (2009). A time-series approach to estimate soil moisture using polarimetric radar data. *IEEE Transactions on Geoscience and Remote Sensing*, 47(8), 2519-2527. Scopus. <https://doi.org/10.1109/TGRS.2009.2014944>

Klopfenstein, T. J., Erickson, G. E., & Berger, L. L. (2013). Maize is a critically important source of food, feed, energy and forage in the USA. *Field Crops Research*, 153, 5-11. <https://doi.org/10.1016/j.fcr.2012.11.006>

Kussul, N., Lavreniuk, M., Skakun, S., & Shelestov, A. (2017). Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. *IEEE Geoscience and Remote Sensing Letters*, 14(5), 778-782. <https://doi.org/10.1109/LGRS.2017.2681128>

Laca, A., Mousia, Z., Díaz, M., Webb, C., & Pandiella, S. S. (2006). Distribution of microbial contamination within cereal grains. *Journal of Food Engineering*, 72(4), 332-338. <https://doi.org/10.1016/j.jfoodeng.2004.12.012>

Lamine, S., Petropoulos, G. P., Brewer, P. A., Bachari, N.-E.-I., Srivastava, P. K., Manevski, K., Kalaitzidis, C., & Macklin, M. G. (2019). Heavy Metal Soil Contamination Detection Using Combined Geochemistry and Field Spectroradiometry in the United Kingdom. *Sensors*, 19(4), Article 4. <https://doi.org/10.3390/s19040762>

Lee, J., Jurkevich, L., Dewaele, P., Wambacq, P., & Oosterlinck, A. (1994). Speckle filtering of synthetic aperture radar images : A Review. *Remote Sensing Reviews*, 8. <https://doi.org/10.1080/02757259409532206>

Lee, J.-S., Wen, J.-H., Ainsworth, T. L., Chen, K.-S., & Chen, A. J. (2009). Improved Sigma Filter for Speckle Filtering of SAR Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 47(1), 202-213. <https://doi.org/10.1109/TGRS.2008.2002881>

Liao, T.-H., Kim, S.-B., Tan, S., Tsang, L., Su, C., & Jackson, T. J. (2016). Multiple Scattering Effects With Cyclical Correction in Active Remote Sensing of Vegetated Surface Using Vector Radiative Transfer Theory. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(4), 1414-1429. <https://doi.org/10.1109/JSTARS.2015.2505638>

- Mandal, D., Kumar, V., Ratha, D., Dey, S., Bhattacharya, A., Lopez-Sanchez, J. M., McNairn, H., & Rao, Y. S. (2020). Dual polarimetric radar vegetation index for crop growth monitoring using sentinel-1 SAR data. *Remote Sensing of Environment*, 247, 111954. <https://doi.org/10.1016/j.rse.2020.111954>
- Meyer, F. (2019). *Spaceborne Synthetic Aperture Radar : Principles, Data Access, and Basic Processing Techniques*. <https://doi.org/10.25966/ez4f-mg98>
- Mulla, D. J. (2013). Twenty five years of remote sensing in precision agriculture : Key advances and remaining knowledge gaps. *Biosystems Engineering*, 114(4), 358-371. <https://doi.org/10.1016/j.biosystemseng.2012.08.009>
- Ndikumana, E., Ho Tong Minh, D., Baghdadi, N., Courault, D., & Hossard, L. (2018). Deep Recurrent Neural Network for Agricultural Classification using multitemporal SAR Sentinel-1 for Camargue, France. *Remote Sensing*, 10(8), Article 8. <https://doi.org/10.3390/rs10081217>
- Ngoune Tandzi, L., & Mutengwa, C. S. (2020). Estimation of Maize (*Zea mays* L.) Yield Per Harvest Area : Appropriate Methods. *Agronomy*, 10(1), Article 1. <https://doi.org/10.3390/agronomy10010029>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., & Cournapeau, D. (2011). Scikit-learn : Machine Learning in Python. *MACHINE LEARNING IN PYTHON*.
- Picard, G., Le Toan, T., & Mattia, F. (2003). Understanding C-band radar backscatter from wheat canopy using a multiple-scattering coherent model. *IEEE Transactions on Geoscience and Remote Sensing*, 41(7), 1583-1591. <https://doi.org/10.1109/TGRS.2003.813353>
- Prasad, A. K., Chai, L., Singh, R. P., & Kafatos, M. (2006). Crop yield estimation model for Iowa using remote sensing and surface parameters. *International Journal of Applied Earth Observation and Geoinformation*, 8(1), 26-33. <https://doi.org/10.1016/j.jag.2005.06.002>
- Pringle, G. (2017, avril 14). *Maize production : Managing critical plant growth stages*. Farmer's Weekly. <https://www.farmersweekly.co.za/crops/field-crops/maize-production-managing-critical-plant-growth-stages/>
- Quarmby, N. A., Milnes, M., Hindle, T. L., & Silleos, N. (1993). The use of multi-temporal NDVI measurements from AVHRR data for crop yield estimation and prediction. *International Journal of Remote Sensing*, 14(2), 199-210. <https://doi.org/10.1080/01431169308904332>
- Ranum, P., Peña-Rosas, J. P., & Garcia-Casal, M. N. (2014). Global maize production, utilization, and consumption. *Annals of the New York Academy of Sciences*, 1312(1), 105-112. <https://doi.org/10.1111/nyas.12396>

scikit-learn.org. (s. d.). *3.1 Cross-validation : Evaluating estimator performance*. Scikit-Learn. Consulté 3 juin 2024, à l'adresse https://scikit-learn/stable/modules/cross_validation.html

SNAP - ESA Sentinel Application Platform v9.0.0 (Version 9.0.0). (2022). [Logiciel]. <https://step.esa.int/main/download/snap-download/>

Su, Q., Lv, J., Fan, J., Zeng, W., Pan, R., Liao, Y., Song, Y., Zhao, C., Qin, Z., & Defourny, P. (2023). Remote Sensing-Based Classification of Winter Irrigation Fields Using the Random Forest Algorithm and GF-1 Data : A Case Study of Jinzhong Basin, North China. *Remote Sensing*, 15, 4599. <https://doi.org/10.3390/rs15184599>

TerraSAR-X - ESA Charter Mapper. (s. d.). Consulté 10 juin 2024, à l'adresse <https://docs.disasterscharter.org/missions/sar/tsx-1/>

Tolera, A., Sundstøl, F., & Said, A. N. (1998a). The effect of stage of maturity on yield and quality of maize grain and stover. *Animal Feed Science and Technology*, 75(2), 157-168. [https://doi.org/10.1016/S0377-8401\(98\)00192-8](https://doi.org/10.1016/S0377-8401(98)00192-8)

Tolera, A., Sundstøl, F., & Said, A. N. (1998b). The effect of stage of maturity on yield and quality of maize grain and stover. *Animal Feed Science and Technology*, 75(2), 157-168. [https://doi.org/10.1016/S0377-8401\(98\)00192-8](https://doi.org/10.1016/S0377-8401(98)00192-8)

Torbick, N., Chowdhury, D., Salas, W., & Qi, J. (2017). Monitoring Rice Agriculture across Myanmar Using Time Series Sentinel-1 Assisted by Landsat-8 and PALSAR-2. *Remote Sensing*, 9(2), Article 2. <https://doi.org/10.3390/rs9020119>

Veloso, A., Mermoz, S., Bouvet, A., Le Toan, T., Planells, M., Dejoux, J.-F., & 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. <https://doi.org/10.1016/j.rse.2017.07.015>

Vreugdenhil, M., Navacchi, C., Bauer-Marschallinger, B., Hahn, S., Steele-Dunne, S., Pfeil, I., Dorigo, W., & Wagner, W. (2020). Sentinel-1 Cross Ratio and Vegetation Optical Depth : A Comparison over Europe. *Remote Sensing*, 12(20), Article 20. <https://doi.org/10.3390/rs12203404>

Vreugdenhil, M., Wagner, W., Bauer-Marschallinger, B., Pfeil, I., Teubner, I., Rüdiger, C., & Strauss, P. (2018). Sensitivity of Sentinel-1 Backscatter to Vegetation Dynamics : An Austrian Case Study. *Remote Sensing*, 10(9), Article 9. <https://doi.org/10.3390/rs10091396>

Wang, C., Wu, J., Zhang, Y., Pan, G., Qi, J., & Salas, W. A. (2009). Characterizing L-Band Scattering of Paddy Rice in Southeast China With Radiative Transfer Model and Multitemporal ALOS/PALSAR Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 47(4), 988-998. <https://doi.org/10.1109/TGRS.2008.2008309>

Xu, J., Meng, J., & Quackenbush, L. J. (2019). Use of remote sensing to predict the optimal harvest date of corn. *Field Crops Research*, 236, 1-13. <https://doi.org/10.1016/j.fcr.2019.03.003>

Zhao, X., Tong, C., Pang, X. M., Wang, Z., Guo, Y., Du, F., & Wu, R. (2011). Functional mapping of ontogeny in flowering plants. *Briefings in bioinformatics*, 13, 317-328. <https://doi.org/10.1093/bib/bbr054>

Zhou, M., & Yang, J. (2023). Delaying or promoting? Manipulation of leaf senescence to improve crop yield and quality. *Planta*, 258(3), 48. <https://doi.org/10.1007/s00425-023-04204-1>

Monitoring the Plant Water Content in Senescent Maize at Parcel Level Using Multi-frequency SAR Data

A Case Study in the Taiyuan Basin, China

Sébastien Saelens

Maize is one of the most widely grown crops globally, playing a pivotal role in agricultural economies and food security. Monitoring the water and dry matter content of maize crops as they reach their final developmental stages is important for ensuring optimal crop yields.

This master thesis explores the potential of SAR data, specifically from TerraSAR-X and Sentinel-1 satellites, in assessing water content and biomass dynamics in senescent maize plants by comparing it with in situ-data from the Taiyuan Basin, China. With support from Chinese partners from the Taiyuan University of Technology, field data on senescent maize was collected in the Taiyuan Basin in August and September 2023. Biophysiological parameters including wet and dry biomass and water content were measured on five selected parcels. These efforts aimed to gather comprehensive in-situ data on the final developmental stages of maize crops.

A temporal analysis of biomass variables was performed to identify discernible patterns and trends in the maturation process of maize. These patterns were then integrated with concurrent SAR data to comprehensively understand the relationships between in-situ observations and remote sensing data. A linear and non-linear correlation analysis between SAR-derived metrics and field measurements were conducted to quantify their intrinsic relationships.

Although the study was not able to establish close correlation between in-situ data and the SAR-derived metrics, it was able to present some unique challenges linked to obtaining in-situ data in the Taiyuan Basin in China, namely the small size of the selected parcels which are often composed of multiple narrow fields and farmed by different owners, leading to high parcel heterogeneity.

These findings highlight the complexities of remote sensing in heterogeneous agricultural landscapes and underscore the need for more refined methods and larger, more uniform study areas. Future research could build on this work to develop more robust techniques for integrating SAR data with ground measurements, ultimately improving the precision of agricultural monitoring and management.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Faculté des bioingénieurs

Croix du Sud, 2 bte L7.05.01, 1348 Louvain-La-Neuve, Belgique | www.uclouvain.be/agro