

Faculté des bioingénieurs

Multi-frequency SAR time series for the detection of seedbed preparation and early-season crop classification

Author	Boland Basile
Supervisor	Defourny Pierre (ELI)
Co-supervisor	Bouchat Jean (ELI)
Readers	Bielders Charles (ELI) Lambot Sébastien (ELI)

Academic Year 2023 - 2024

Master thesis presented in fulfillment of the requirements for the degree of
Bioengineer: Environmental Science and Technology

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to Pierre Defourny, my supervisor. His guidance and support have been very important in my journey through the field of geomatics, a discipline I have grown to love during my final years as a student. His vast experience and profound knowledge consistently pushed me to improve and strive for excellence.

Secondly, I wish to extend my sincere thanks to Jean Bouchat, my co-supervisor and final-year PhD candidate in the Geomatics Research team of UCLouvain. His patience and experienced guidance have been essential as I faced the complexities of scientific research. Despite being in the final stages of his own PhD thesis, he always found time to share his insights and provide invaluable advice to help me improve.

I am also deeply grateful to Tom Kenda, another dedicated PhD candidate of the Geomatics Research team. His considerable efforts and the time he invested in helping me gather information and providing constructive feedback were crucial to the progress of my work.

On a personal note, I would also like to thank my parents and brother, whose unconditional support has been a pillar of strength throughout my years as a student in Bioengineering. I am also profoundly grateful to my girlfriend, Laëtitia, who, despite being a final-year medical student herself, always found ways to encourage and support me. Finally, I want to thank my friends, Julien and Mattias, for their support and collaboration over the past few months.

CONTENTS

List of Figures	iii
List of Tables	ix
List of Acronyms	x
Introduction	1
1 Literature review	2
1.1 Agricultural challenges	2
1.1.1 Croplands in Belgium	2
1.1.2 Soil management practices: conventional and conservation agriculture	4
1.1.3 Sowing practices: a matter of precision	6
1.2 Evolution and improvements in agricultural remote sensing technologies . .	7
1.2.1 SAR remote sensing: basic principles applied to soil monitoring . .	7
1.2.2 Polarimetric SAR data to estimate soil parameters	8
1.3 Change detection and monitoring of soil preparation activities using SAR .	10
1.3.1 First tillage change detection applications	10
1.3.2 Radar and optical imagery: a hybrid method	10
1.3.3 Coherent change detection (CCD) for agricultural activities	11
1.4 Algorithms for crop classification	13
1.4.1 K-Nearest Neighbours: simple yet powerful	13
1.4.2 Support Vector Machine: high-dimensional spaces	14
1.4.3 Random Forest: a decision tree ensemble	15
1.4.4 Comparison between KNN, SVM and RF: pros and cons	16
1.4.5 Accuracy metrics for classification	17
1.5 Conclusion on current scientific knowledge	18
2 Objectives	19
3 Materials and methods	20
3.1 Study area: Beglian’s silty region	20
3.1.1 Soil characteristics: fertility in the silty region	21
3.1.2 2021 agricultural season: a cold spring and wet summer	22
3.2 Available data	24
3.2.1 In-situ data collection	24
3.2.2 SAR imagery: X-band and C-band	25

3.2.3	Precipitation data	27
3.3	Data processing	27
3.3.1	Field polygons	27
3.3.2	Zonal statistics, indices and coherence	27
3.3.3	Time series	29
3.4	Seedbed preparation date detection	29
3.4.1	Coherence time series	29
3.4.2	In-situ data	30
3.5	Early season crop type classification	31
3.5.1	Feature extraction	31
3.5.2	Random Forest classification	31
4	Results	33
4.1	Time series: a first visual analysis	33
4.1.1	TerraSAR-X time series	33
4.1.2	Sentinel-1 time series	34
4.2	Seedbed preparation detection	35
4.2.1	Sample fields: TerraSAR-x & Sentinel-1	35
4.2.2	Sentinel-1 footprint	39
4.2.3	Validation	41
4.3	Early-season crop classification	42
4.3.1	TerraSAR-X	43
4.3.2	Sentinel-1	47
4.3.3	Sentinel-1 reduced frequency for comparison purposes	51
4.3.4	TerraSAR-X and Sentinel-1 data combination	52
4.3.5	Classification accuracy: summary	55
5	Discussion	56
5.1	Seedbed preparation detection with InSAR coherence time series	56
5.2	Early-season crop classification	57
5.2.1	Impact of the minimum pixel threshold on classification results	57
5.2.2	X-band and C-band: respective contributions and comparison	58
5.2.3	Combination of X-band and C-band: multi-frequency as a solution?	60
5.2.4	Importance of first acquisitions for early classification	60
5.3	Limitations and perspectives	61
6	Conclusion	63
7	Bibliography	i
8	Appendices	vii
A	Materials and methods	vii
B	Results	viii

LIST OF FIGURES

1.1	Share of agriculture surface in land use in Belgium in 2022 (Statbel, 2023).	3
1.2	Different tools commonly used in conventional farming soil management practices (Bertin, 2023)	5
1.3	Radar reflection from smooth, moderately rough, and very rough surfaces, highlighting specular and diffuse reflection principles in radar backscattering (Sabery et al., 2021).	8
1.4	Visual representation of the three coherence jump detection methods where a jump occurs at T_i . (a) The mean shift method with a window size (d) of 9, where the difference between the smoothed values is evaluated by the red arrow, (b) the linear regression method of window size (d) of 4, where the raw value T_i is compared to the preceded smoothed value and (c) the two means method with a window size (d) of 8 comparing the null (H_0) and alternative hypothesis (H_1) (De Vroey et al., 2021).	13
1.5	Schematic 2D representation of the KNN classification process with two classes. (a) A 1-NN decision rule where the new data point B is assigned to the red square class and (b) a k -NN decision rule where $k = 3$ and B is assigned to the blue circles class.	14
1.6	Schematic visualisation of the main SVM characteristics (Mathworks, n.d.).	15
1.7	Schematic visualisation of the process of a Random Forest algorithm classification (Gunay, 2023).	16
3.1	Map of agricultural regions of Belgium with location of the area of interest (AOI) in the Walloon region near the cities of Gembloux, Walhain and Perwez at the intersection of TerraSAR-X and Sentinel-1 footprints in this area. The AOI mainly covers the silty region but also presents small bits in the Condroz and the sandy silty region.	21
3.2	Share of total area used for agriculture by agricultural regions in 2022. The silty region is represented at the top of the chart and the vertical red line is showing the mean value for the Walloon region (SPW Agriculture, n.d.).	21
3.3	Average daily temperature in Uccle for the 2021 spring season. The black line indicates the normal average temperature between 1991 and 2020. Red and blue bars respectively indicate higher or lower anomalies (IRM, 2021).	22

3.4	Daily and cumulative precipitation in Uccle for June 2021 (a) and July 2021 (b). Blue bars are representing daily observations with darkness depending on precipitation intensity. The dashed line shows the cumulative average between 1991 and 2020 while the full line shows cumulative observations (IRM, 2021).	23
3.5	Map of visited fields (yellow) during the 2021 field campaign highlighting the campaign itinerary (red).	24
4.1	Mean evolution of the TerraSAR-X Cross-polarization Ratio (CR) and hourly precipitation through the 2021 spring season for all parcels of each studied culture (beet, maize, potato).	34
4.2	Mean evolution of the Sentinel-1 Cross-polarization Ratio (CR) and hourly precipitation through the 2021 spring season for all parcels of each studied culture (beet, maize, potato).	35
4.3	Mean evolution and associated standard deviation of the TerraSAR-X VV (green) & VH (red) backscatter (σ_{VV}^0 & σ_{VH}^0) for all sample fields that share the same sowing period (vertical orange stripe) for each studied crop type.	36
4.4	Mean evolution and associated standard deviation of the TerraSAR-X VV (green) & VH (red) backscatter (σ_{VV}^0 & σ_{VH}^0) and VV (blue) & VH (purple) coherence for all field campaign beet fields that share the same sowing period (vertical orange stripe).	37
4.5	Mean evolution and associated standard deviation of the TerraSAR-X VV (green) & VH (red) backscatter (σ_{VV}^0 & σ_{VH}^0) and VV (blue) & VH (purple) coherence for all field campaign maize fields that share the same sowing period (vertical orange stripe).	38
4.6	Mean evolution and associated standard deviation of the TerraSAR-X VV (green) & VH (red) backscatter (σ_{VV}^0 & σ_{VH}^0) and VV (blue) & VH (purple) coherence for all field campaign potato fields that share the same sowing period (vertical orange stripe).	38
4.7	Spatial distribution map of the seedbed preparation date detected on the Sentinel-1 footprint for beet fields. The color intensity indicates the detected date with a stronger intensity being related to later dates. The bar plot shows the temporal distribution in number of fields, keeping the same color scheme.	39
4.8	Spatial distribution map of the seedbed preparation date detected on the Sentinel-1 footprint for maize fields. The color intensity indicates the detected date with a stronger intensity being related to later dates. The bar plot shows the temporal distribution in number of fields, keeping the same color scheme.	40
4.9	Spatial distribution map of the seedbed preparation date detected on the Sentinel-1 footprint for potato fields. The color intensity indicates the detected date with a stronger intensity being related to later dates. The bar plot shows the temporal distribution in number of fields, keeping the same color scheme.	41
4.10	Histogram of beet sowing dates observed in the IRBAB network for 162 fields (BCGMS, 2021).	42

4.11	Evolution of F1 score for Random Forest classification of beet, maize, and potato crops on the area of interest with cumulative addition of TerraSAR-X acquisition data, while maintaining a minimum threshold of 50 pixels of image per field.	44
4.12	Evolution of F1 score for Random Forest classification of beet, maize, and potato crops on the area of interest with cumulative addition of TerraSAR-X acquisition data, while maintaining a minimum threshold of 100 pixels of image per field.	44
4.13	Evolution of F1 score for Random Forest classification of beet, maize, and potato crops on the area of interest with cumulative addition of TerraSAR-X acquisition data, while maintaining a minimum threshold of 150 pixels of image per field.	45
4.14	Top ten of most important features for the Random Forest classification of beet, maize, and potato crops using TerraSAR-X data on the area of interest, while maintaining a minimum threshold of 50 pixels of image per field.	46
4.15	F1 score evolutions for the Random Forest classifications of beet, maize, and potato crops on the area of interest with cumulative addition of TerraSAR-X acquisition data. Each plot shows classification performance where the input data progressively exclude the earliest acquisitions. A minimum threshold of 50 pixels of image per field is applied.	46
4.16	Evolution of F1 score for Random Forest classification of beet, maize, and potato crops on the area of interest with cumulative addition of Sentinel-1 acquisition data, while maintaining a minimum threshold of 50 pixels of image per field.	47
4.17	Evolution of F1 score for Random Forest classification of beet, maize, and potato crops on the area of interest with cumulative addition of Sentinel-1 acquisition data, while maintaining a minimum threshold of 100 pixels of image per field.	48
4.18	Evolution of F1 score for Random Forest classification of beet, maize, and potato crops on the area of interest with cumulative addition of Sentinel-1 acquisition data, while maintaining a minimum threshold of 150 pixels of image per field.	48
4.19	Top ten of most important features for the Random Forest classification of beet, maize, and potato crops using Sentinel-1 data on the area of interest, while maintaining a minimum threshold of 50 pixels of image per field. . .	49
4.20	F1 score evolutions for the Random Forest classifications of beet, maize, and potato crops on the area of interest with cumulative addition of Sentinel-1 acquisition data. Each plot shows classification performance where the input data progressively exclude the earliest acquisitions. A minimum threshold of 50 pixels of image per field is applied.	50
4.21	Evolution of F1 score for Random Forest classification of beet, maize, and potato crops on the complete Sentinel-1 footprint with cumulative addition of Sentinel-1 acquisition data, while maintaining a minimum threshold of 50 pixels of image per field.	51

4.22	Evolution of F1 score for Random Forest classification of beet, maize, and potato crops on the area of interest with cumulative addition of a reduced amount of 10 Sentinel-1 acquisitions data, while maintaining a minimum threshold of 50 pixels of image per field.	51
4.23	Evolution of F1 score for Random Forest classification of beet, maize, and potato crops on the area of interest with cumulative addition of a combination of TerraSAR-X and Sentinel-1 acquisitions data, while maintaining a minimum threshold of 50 pixels of image per field.	52
4.24	Top ten of most important features for the Random Forest classification of beet, maize, and potato crops using Sentinel-1 and TerraSAR-X data on the area of interest, while maintaining a minimum threshold of 50 pixels of image per field.	53
4.25	F1 score evolutions for the Random Forest classifications of beet, maize, and potato crops on the area of interest with cumulative addition of TerraSAR-X and Sentinel-1 acquisitions data. Each plot shows classification performance where the input data progressively exclude the earliest acquisitions. A minimum threshold of 50 pixels of image per field is applied.	53
4.26	Four first confusion matrices derived from the Random Forest classification of beet, maize, and potato crops on the area of interest with cumulative addition of a combination of TerraSAR-X and Sentinel-1 acquisitions data, while maintaining a minimum threshold of 50 pixels of image per field. . .	54
8.1	Daily and cumulative precipitation in Uccle for the 2021 spring season. Blue bars are representing daily observations with darkness depending on precipitation intensity. The dashed line shows the cumulative average between 1991 and 2020 while the full line shows cumulative observations (IRM, 2021).	vii
8.2	Mean evolution of the TerraSAR-X VV backscatter (σ_{VV}^0) and hourly precipitation through the 2021 spring season for all parcels of each studied culture (beet, maize, potato).	viii
8.3	Mean evolution of the TerraSAR-X VH backscatter (σ_{VH}^0) and hourly precipitation through the 2021 spring season for all parcels of each studied culture (beet, maize, potato).	viii
8.4	Mean evolution of the TerraSAR-X DpRVI and hourly precipitation through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).	ix
8.5	Mean evolution of the TerraSAR-X VV backscatter (σ_{VV}^0) with standard deviation buffer through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).	ix
8.6	Mean evolution of the TerraSAR-X VH backscatter (σ_{VH}^0) with standard deviation buffer through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).	x
8.7	Mean evolution of the TerraSAR-X CR with standard deviation buffer through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).	x
8.8	Mean evolution of the TerraSAR-X DpRVI with standard deviation buffer through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).	xi

8.9	Mean evolution of the Sentinel-1 VV backscatter (σ_{VV}^0) and hourly precipitation through the 2021 spring season for all parcels of each studied culture (beet, maize, potato)).	xi
8.10	Mean evolution of the Sentinel-1 VH backscatter (σ_{VH}^0) and hourly precipitation through the 2021 spring season for all parcels of each studied culture (beet, maize, potato).	xii
8.11	Mean evolution of the Sentinel-1 DpRVI and hourly precipitation through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).	xii
8.12	Mean evolution of the Sentinel-1 VV coherence and hourly precipitation through the 2021 spring season for all parcels of each studied culture (beet, maize, potato).	xiii
8.13	Mean evolution of the Sentinel-1 VH coherence and hourly precipitation through the 2021 spring season for all parcels of each studied culture (beet, maize, potato).	xiii
8.14	Mean evolution of the Sentinel-1 VV backscatter (σ_{VV}^0) with standard deviation buffer through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).	xiv
8.15	Mean evolution of the Sentinel-1 VH backscatter (σ_{VH}^0) with standard deviation buffer through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).	xiv
8.16	Mean evolution of the Sentinel-1 CR with standard deviation buffer through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).	xv
8.17	Mean evolution of the Sentinel-1 DpRVI with standard deviation buffer through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).	xv
8.18	Mean evolution of the Sentinel-1 VV coherence with standard deviation buffer through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).	xvi
8.19	Mean evolution of the Sentinel-1 VH coherence with standard deviation buffer through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).	xvi
8.20	Top ten of most important features for the Random Forest classification of beet, maize, and potato crops using TerraSAR-X data on the area of interest, while maintaining a minimum threshold of 100 pixels of image per field.	xvii
8.21	Top ten of most important features for the Random Forest classification of beet, maize, and potato crops using TerraSAR-X data on the area of interest, while maintaining a minimum threshold of 150 pixels of image per field.	xvii
8.22	Top ten of most important features for the Random Forest classification of beet, maize, and potato crops using Sentinel-1 data on the area of interest, while maintaining a minimum threshold of 100 pixels of image per field. . .	xviii
8.23	Top ten of most important features for the Random Forest classification of beet, maize, and potato crops using Sentinel-1 data on the area of interest, while maintaining a minimum threshold of 150 pixels of image per field. . .	xviii

- 8.24 Last six confusion matrices derived from the Random Forest classification of beet, maize, and potato crops on the area of interest with incremental addition of a combination of TerraSAR-X and Sentinel-1 acquisitions data, while maintaining a minimum threshold of 50 pixels of image per field. . . . xix

LIST OF TABLES

1.1	Crop calendar of the main spring crops encountered in Wallonia in 2011 (Biielders et al., 2011)	4
1.2	Main SAR frequency bands used in remote sensing and their associated frequency and wavelength range (AMS, 2022).	8
1.3	Thresholding of VH backscatter changes at High (ΔVH_{HR}) and Medium Resolution (ΔVH_{MR}) with th_{XX} being the threshold value at a certain resolution (Satalino et al., 2018).	11
1.4	Advantages and disadvantages of K-Nearest Neighbour, Support Vector Machine and Random Forest classification algorithms (Boateng et al., 2020).	17
3.1	Number of fields being sown for each crop type (beet, potato and maize) between each visit during the 2021 field campaign. The period refers to the time between two field visits where soil management status transitions from "deep ploughing" to "prepared seedbed". The most dominant cluster for each crop type is highlighted with a different color. The date format is set as mm-dd.	30
4.1	Statistics of sowing dates (2021) for beet fields in the Silty region and the Condroz (BCGMS, 2021).	41
4.2	Minimum surface covered in ha corresponding to every minimum pixel value threshold per field for TerraSAR-X and Sentinel-1.	43
4.3	Class distribution for each crop type (beet, maize, potato) on the area of interest depending on the minimum image pixel threshold set by field for TerraSAR-X.	43
4.4	Class distribution for each crop type (beet, maize, potato) on the area of interest depending on the minimum image pixel threshold set by field for Sentinel-1.	47
4.5	Earliest acquisition date at which the F1 score of each Random Forest classification process for beet, maize and potato fields on the area of interest reached 0.8. Different scenarios are considered for TerraSAR-X (TSX) and Sentinel-1 (S1) with varying thresholds of minimum pixels covering each field. Color intensity related to the acquisition date is darker when the score is reached earlier in the season and lighter when the score is reached later in the season.	55

LIST OF ACRONYMS

AEIM	Advanced Integral Equation Model
BMP	Beneficial Management Practices
CA	Conservation Agriculture
CAP	Common Agricultural Policy
CCD	Coherent Change Detection
CR	Cross-polarization Ratio
DEM	Digital Elevation Model
DN	Digital Numbers
DpRVI	Dual-polarization Radar Vegetation Index
EEC	Enhanced Ellipsoid Corrected
EU	European Union
GSD	Ground Sampling Distance
HH	Horizontal send and horizontal receive polarization
IEM	Integral Equation Model
InSAR	Interferometric Synthetic Aperture Radar
IW	Interferometric Wide
KNN	K-Nearest Neighbours
LPIS	Land Parcel Identification System
MDI	Mean Decrease in Impurity
OA	Overall Accuracy
px	Pixel(s)
RE	Radiometric Enhancement
RMI	Royal Meteorological Institute
rmsh	Root mean square height
RF	Random Forest
S1	Sentinel-1
SAR	Synthetic Aperture Radar
SDG	Sustainable Development Goals
SLC	Single-Look Complex
SM	Stripmap Mode
SMOTE	Synthetic Minority Over-sampling Technique
SVM	Support Vector Machine
TSX	TerraSAR-X
UAA	Used Agricultural Area
VH	Vertical send and horizontal receive polarization
VV	Vertical send and vertical receive polarization

INTRODUCTION

The 21st century is slowly progressing towards the construction of a more sustainable future thanks to increased awareness of environmental issues, advancements in green technology, and a growing commitment from both individuals and organizations to reduce their carbon footprint. Central to this movement are the United Nations' Sustainable Development Goals (SDGs), which provide a comprehensive framework for achieving a sustainable future by 2030 (United Nations, n.d.).

One of the most pressing challenges on this journey towards sustainability is the significant pressure on agricultural and natural resources, driven by increasing population and consumption in recent years (Foley et al., 2011). In 2022, 2.4 billion people did not have access to nutritious, safe and sufficient food all year round while our current agricultural systems are degrading water, land, biodiversity and climate on a global scale (FAO, 2023; Foley et al., 2011). It is also expected that almost 600 million people will be chronically undernourished in 2030, an increasing number during the past years that are closely related to the recent COVID pandemic and still ongoing Ukrainian war (FAO, 2023). Agriculture plays a crucial role in the sustainable transition, particularly in achieving SDG 2 (Zero Hunger) and SDG 12 (Responsible Consumption and Production). It is therefore clear that the agricultural sector needs to be well understood and managed in order to fulfill the world's consumption needs with respect to the Earth's limited resources.

Remote sensing offers the capability to gather information for monitoring and mapping crops across spatial and temporal scales that are challenging to achieve through in situ methods (Bahrami et al., 2022). Diverse sensor technologies and observation platforms have been developed to monitor Earth's surface. Optical remote sensing for crop monitoring was firstly used in the early 1970s with the first Landsat satellite (Becker-Reshef et al., 2010). With spatial, spectral and temporal resolution getting better through the years, these types of sensors as well as radar sensors have since been used extensively for agricultural purposes (McNairn et al., 2014).

LITERATURE REVIEW

This chapter presents a summary of current scientific knowledge related to the thesis topic, guiding the formulation of its scientific contribution, experimental protocol, and result interpretation.

1.1 Agricultural challenges

With sustainable development at the center of most recent debates, agriculture faces a dual challenge: reducing agricultural intensity to minimize its environmental impact while maintaining sufficient yields to meet the world's consumption needs. One key aspect of addressing this challenge lies in accurate crop production forecasts, which play a vital role in ensuring food security and anticipating production shortfalls (Jayne and Rashid, 2010). Early and precise yield projections not only limit food price volatility but also increase resilience against food insecurity (Jayne and Rashid, 2010).

Moreover, sustainable agriculture emphasises on effective soil management practices. For instance, different intensive tillage techniques have strong influence on the soil biology and crop production, its monitoring should therefore become of great importance in agricultural sciences (Satalino et al., 2018). Tillage operations affect multiple land dynamics such as soil erosion, run-off and infiltration, carbon sequestration and water and carbon dioxide exchanges (Foley et al., 2011).

1.1.1 Croplands in Belgium

All of the above holds particular significance in Belgium, where over 40% of the country's surface is dedicated to agricultural crop fields (Statbel, 2021) with an even higher percentage in the silty, sandy silty and condroz regions for Wallonia as shown on Figure 1.1.

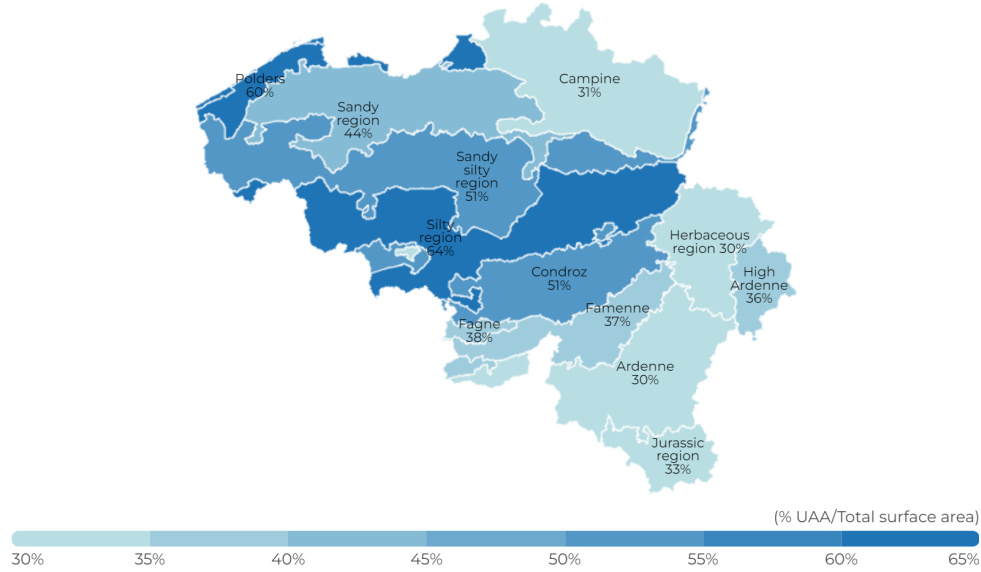


Figure 1.1: Share of agriculture surface in land use in Belgium in 2022 (Statbel, 2023).

Belgium faces a challenge concerning the peri-urban nature of its agricultural areas with competition from housing, industrial zones, and infrastructure development (Diplomatie Belgium, 2023). This might lead to geographical proximity between rural and agricultural areas exposing housings to higher risks of flooding and mudflows if agricultural lands are not well managed.

Belgium also operates within the framework of the Common Agricultural Policy (CAP) of the European Union (EU), aligning its agricultural practices with broader EU strategies and objectives. Some of the CAP 2023-2027 objectives are: climate change action, environmental care, landscape and biodiversity preservation and the forming of vibrant rural areas (European Comission, n.d.).

With the potential negative effects on the environment and biodiversity of intensive soil tillage practices and the high agricultural density of the country, Belgium has to put double effort in managing and monitoring soil preparation techniques.

In Belgium, a diverse range of crops are cultivated, but some of the most common spring crop types found in Wallonia are sugar beets, potatoes, and maize, as identified by the 2021 Land Parcel Identification System (see Section 3.2). The following study will concentrate on these specific species. Table 1.1 illustrates the crop calendar for selected spring crop types during a standard year.

Table 1.1: Crop calendar of the main spring crops encountered in Wallonia in 2011 (Bielders et al., 2011)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Beet												
Potato												
Maize												

Sowing

Mid-season

Harvest

1.1.2 Soil management practices: conventional and conservation agriculture

The two following sections (1.1.2 & 1.1.3) are widely based on the "LBRAI2106 - Phytotechnie" course given by Pierre Bertin at the faculty of Bioscience Engineering, UCLouvain (Bertin, 2023).

When discussing soil management practices, two main groups of methods can be identified: conventional tillage practices and alternative methods. Conventional methods primarily involve various tillage techniques, while alternative methods include shallow ploughing and reduced tillage techniques. Each approach has its own set of advantages and disadvantages.

Conventional tillage techniques

The conventional approach can be subdivided in four methods, each one using different tools. Firstly, the deep ploughing (Figure 1.2a) method which involves burying organic amendments and residues, destroying weeds, and loosening the soil to a depth of 20-30 cm using a plow. This is the only method that is overturning the upper soil layers. Secondly, the light cultivator tool (Figure 1.2b) smoothens the furrows left by the ploughing, working at a depth of 5 to 15 cm to homogenize the soil and destroy possible germinated weeds. Thirdly, the rotative harrow (Figure 1.2c) is an animated tool with a vertical rotation axis that mixes soil at a surface level of 10 to 15 cm, preparing seedbed. This tool is suitable for wetter soils but increases risks of soil compaction. Finally, the clod breaker roll (Figure 1.2d) can be used to further homogenize the soil, flatten larger clumps and ensure proper seed-soil contact.

Advantages of conventional tilling activities include the effective burial of organic amendments like lime, manure, crop residues, and pesticide residues, thereby incorporating them into the soil. Additionally, they help in the creation of a fine soil structure through exposure to frost action, enhance soil aeration via mineralization, and facilitate drainage due to fast spring warming.

However, these methods also harbor several drawbacks, including deep soil compaction and the formation of a plow pan. They can bring long-lived weed seeds to the surface, which may germinate years later. Moreover, the high energy costs and carbon footprint associated with ploughing pose significant concerns. Excessive mineralization can lead to the loss of organic matter and biological activity, further exacerbating soil erosion and nitrogen loss.



Figure 1.2: Different tools commonly used in conventional farming soil management practices (Bertin, 2023)

Soil management in conservation agriculture

Conservation agriculture is mainly defined through three important principles: limited tillage, winter cover crops and crop rotation and diversification. Several alternatives to conventional ploughing methods exist. Two main sub-groups can be identified: shallow ploughing and beneficial management practices (BMPs). Shallow ploughing involves using narrower plows with smaller shares to disturb the soil less deeply, thereby reducing the creation of compacted plow pans. Beneficial management practices involve minimizing or eliminating ploughing. This approach is gaining popularity in both conventional

and conservation agriculture, with approximately 50% of wheat and rapeseed fields not being ploughed conventionally in Wallonia. Three main categories of BMPs can be fixed. Firstly, pseudo-ploughing, which mimics the effects of deep ploughing (> 15 cm) without soil overturning. This technique helps to alleviate soil compaction without fully disturbing the soil layers. Secondly, a light cultivator tool (< 15 cm) can be used to perform very light soil work, implying the soil is not deeply compacted. Finally, the direct sowing method consists in the planting of seeds into untilled soil.

Advantages of BMPs include time savings during crucial periods of the year, fuel economy with potential savings of up to 30% with reduced tillage and up to 60% with no-till practices, soil structure improvement leading to reduced compaction and erosion, and enhancement of biological activity, especially when combined with vegetation cover.

Nevertheless, BMPs also present several notable disadvantages, including the requirement for more observation and expertise, challenges in weed management, longer crop rotations, and higher risks of fusarium and other diseases transmitted by non-buried crop residues.

1.1.3 Sowing practices: a matter of precision

Soil management is crucial for a successful sowing. Clods on the soil surface should be of approximately 5 cm to better resist to compaction while allowing regular germination. The seedbed should be a fine layer of soil overlaying an uncompacted soil for good drainage. Two mistakes to avoid are seeding too deep and excessively fine soil preparation (risk of crusting). Seeding depth should also be adapted according to soil moisture.

Different sowing machines exist, used for different purposes and crop types. Precision seeders allow precise control of the distance between each seed on the field. This is crucial for plants with significant individual growth that don't tolerate close competition from neighbouring plants. Examples include root crops such as beets, chicory or spinach. Standard seeders, on the other hand, do not offer the control of seed spacing. Seed rates are calculated per hectare, but seed distribution may not be precise. This seeding method is more suitable for crops like cereals that tolerate competition.

Beet seeds are small and fragile, requiring stronger attention to meticulous soil preparation and strict control of germination to avoid loss, with an emergence rate of 70 to 80%. Choosing the sowing period involves a compromise: sowing too early in the season increases the risk of early vernalization, while sowing too late reduces sugar yields. Due to climate change, sowing, which occurs in mid-April, can sometimes already take place in March. Seeds for potato and maize (mainly silage maize in Wallonia) are less fragile and require less attention regarding seed quality. Rougher soil preparation is preferable,

to limit evaporation. Sowing generally takes place later in the season, around mid-April or early May. Let's note that focus is here set on the details concerning selected spring crop types for the study, but all of above mentioned techniques can also be applied for winter crops in Wallonia.

1.2 Evolution and improvements in agricultural remote sensing technologies

1.2.1 SAR remote sensing: basic principles applied to soil monitoring

While presence of clouds remain a major drawback for optical sensors, radar remote sensing has the ability to acquire day-and-night images under almost all-weather conditions due to the use of an active radiation source in the microwave frequency range (Van Tricht et al., 2018). The Synthetic Aperture Radar (SAR) technology is widely spread in radar remote sensing and uses the motion of the instrument in order to attain satisfactory ground resolution. As the SAR device moves, multiple recordings from different antenna positions are processed to form a synthetic aperture, enabling the creation of detailed images not easily achievable with a fixed physical antenna, which should be much larger in size (Campbell and Wynne, 2011).

The recent availability of satellite imagery data acquired with large spatial and temporal resolutions opens new perspectives for identifying soil management activities over crop fields, which can happen in space and time. SAR imagery is also very well known for its abilities to detect surface roughness and soil dielectric properties, being linked to soil moisture (Hadria et al., 2009; McNairn et al., 1996). The soil roughness influences the backscattering type and therefore the intensity of the returned signal. More specifically, C-band SAR response is very sensitive to changes in structure during seed and fruit development (McNairn et al., 2009).

Figure 1.3 highlights how different surface roughness types will impact the signal returned to the sensor. A smooth surface will reflect a large amount of the energy in the specular direction, letting little to no signal reach the antenna. On the other hand, a rough surface will diffuse the incident wave in all directions (Baghdadi et al., 2012b). This effect is directly related to the relationship between the wavelength of the signal and the roughness level, as described by the Rayleigh criterion. According to this criterion, a smooth surface is characterized by a phase difference smaller than one quarter of the signal wavelength (Defourny, 2022). Given that the signal must travel a two-way path to reach the sensor again after being scattered, the mathematical expression of a smooth

surface stands as described as follows:

$$rmsh < \frac{\lambda}{8 \cos \theta_i} \quad (1.1)$$

where $rmsh$ represents the surface roughness, defined as the root mean square height relative to a perfectly smooth surface, λ the signal wavelength and θ_i the angle of incidence (EO College, 2023).

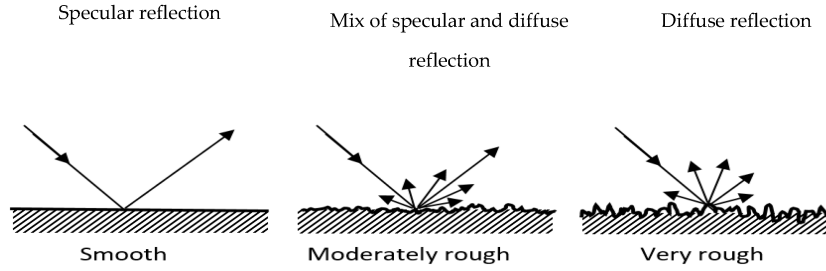


Figure 1.3: Radar reflection from smooth, moderately rough, and very rough surfaces, highlighting specular and diffuse reflection principles in radar backscattering (Sabery et al., 2021).

1.2.2 Polarimetric SAR data to estimate soil parameters

As mentioned in the above section, radar sensors send pulses in the microwave frequency. This microwave spectrum can be separated in different bands depending on their frequency range (or associated wavelength range). Some of the main radar frequency bands are described in Table 1.2.

Table 1.2: Main SAR frequency bands used in remote sensing and their associated frequency and wavelength range (AMS, 2022).

Frequency band	Frequency range (GHz)	Wavelength range (cm)
L	1-2	15-30
S	2-4	7.5-15
C	4-8	3.75-7.5
X	8-12	2.5-3.75

L-band SAR data has been widely used for soil properties characterisation under canopy (Mattia et al., 2009) but with development of observation frequency and availability of C-band SAR imagery, multiple studies have been conducted to extract soil moisture information from mono- or multi-polarization C-band SAR data (Álvarez-Mozos et al., 2006; Baghdadi et al., 2002, 2006; Paloscia et al., 2008). C-band measurements are impacted by factors such as soil surface roughness and humidity (Chomé et al., 2016).

Changes in backscattering coefficients serve as indicators of soil surface activities like ploughing and sowing, fluctuations in soil moisture levels, and the growth or removal of crop canopy (Lemoine et al., 2002; McNairn et al., 1998). Tilling the agricultural soils can therefore either increase roughness, with an initial tillage pass on untilled land, or decrease roughness, as seen with a secondary or tertiary tillage aimed at preparing the seedbed (Robertson et al., 2023).

Past studies had shown that the most optimal way of estimating soil moisture and surface roughness for one polarization and one radar wavelength are obtained using SAR images acquired with two different incidence angles (20° and 40° for example) (Baghdadi et al., 2006; Srivastava et al., 2003). Nowadays, the great majority of SAR sensors are polarimetric, meaning they can capture data for multiple polarizations simultaneously.

As also previously introduced through the Rayleigh criterion, roughness in microwave modeling is expressed as the statistical variation of soil surface indicated by the root mean square height ($rmsh$) and correlation length (l) (Robertson et al., 2023). The inversion approach, characterized by a backward working methodology that starts from observed effects or data to deduce underlying causes or processes, has frequently been used in the estimation of soil parameters (Baghdadi et al., 2012a; Satalino et al., 2002). This method can be applied using physical or statistical models to estimate soil parameters from SAR images.

A commonly used physical model to estimate soil moisture and surface roughness is the Integral Equation Model (IEM) with a validity domain in C-band close to common range of roughness values encountered for agricultural surfaces: $k rmsh \leq 3$ (with $rmsh$ being the root mean square height and k the radar wave number $\cong 1.11 \text{ cm}^{-1}$ for a 5.3 GHz frequency) (Fung et al., 1992). Typical $rmsh$ values for bare soils in agriculture range from 0.5 to 4 cm (Baghdadi et al., 2012a).

In this continuity and knowing that the estimation of such variables is a complex and nonlinear process, artificial neural networks have been applied to the process by using simulated radar backscattering coefficients through the IEM. With addition of a priori information, decent results could be obtained reaching an accuracy of 0.5 cm (Root Mean Square Error) for surface roughness for values of $rmsh$ lower than 2 cm. For higher $rmsh$ values, the neural networks underestimate the surface roughness (Baghdadi et al., 2012a). Although the IEM is a commonly used and most adequate model for soil parameter estimation, several other models exist for these purposes including the Advanced Integral Equation Model (AIEM), the Oh model, and the Dubois model (Choker et al., 2017).

1.3 Change detection and monitoring of soil preparation activities using SAR

When farmers till, sow and harvest their fields, soil is disturbed and this disturbance can be measured by SARs (McNairn et al., 2022). To fully understand and accurately detect seedbed preparation, it is important to analyze the changes in soil roughness within a specific crop field. Furthermore, with new European policies such as the CAP, which is pushing towards principles of Conservation Agriculture (CA), farmers are encouraged to adopt BMPs such as little to no tillage (Robertson et al., 2023). Satellites are one way to monitor BMP adoption.

Most of current SAR satellites typically provide a revisit period between image acquisitions ranging from days to weeks (Robertson et al., 2023). Satellite imagery can therefore offer the ability to capture changes in field conditions throughout the busy harvesting and seeding periods (McNairn et al., 1996).

1.3.1 First tillage change detection applications

Most researchers investigating the monitoring of tillage activities have primarily concentrated on recovering the root mean square backscatter intensity (McNairn et al., 1996), identifying changes in backscattering caused by tillage (McNairn et al., 1998), or incorporating SAR images into classifiers (Azzari et al., 2019; Luotamo et al., 2022). Use of different polarization implements (HH for horizontal send and horizontal receive, VV for vertical send and vertical receive and VH for vertical send and horizontal receive) allows to retrieve different signals and thus extract different information. Although change in HH over time on multiple images could be detected and associated with primary tillage (McNairn et al., 1998), presence of crop residue also affected scattering. This is why researchers proposed to classify tillage using both SAR and optical imagery. This highlights the challenge and complexity of using SAR backscatter alone (Robertson et al., 2023).

1.3.2 Radar and optical imagery: a hybrid method

Combining radar (e.g. Sentinel-1 VH) and optical imagery such as Landsat 5,7 and 8 (Azzari et al., 2019) or Sentinel-2 (Satalino et al., 2018) allows to acquire additional information on vegetation or crop residue presence on the field by using different spectral indices such as NDVI, to only name one.

Some research projects (Satalino et al., 2018) focus on using optical imagery to exclude bare or scarcely vegetated areas (over a certain NDVI value) and then using VH backscatter at two different resolutions to analyse roughness changes and identify the reason of change based on a particular thresholding pattern detailed in Table 1.3. Other

authors (Luotamo et al., 2022) used Sentinel-1 VV and VH intensity and two optical indices derived from Sentinel-2 data to extract and classify different tillage practices or crop cover types using different classification techniques, with Random Forest (RF) achieving the best overall accuracy with 70% for the six defined classes.

Table 1.3: Thresholding of VH backscatter changes at High (ΔVH_{HR}) and Medium Resolution (ΔVH_{MR}) with th_{XX} being the threshold value at a certain resolution (Satalino et al., 2018).

Thresholding	Reason of change
$-th_{HR} < \Delta VH_{HR} < th_{HR}$	No change
$\Delta VH_{HR} > th_{HR}$ and $-th_{MR} < \Delta VH_{MR} < th_{MR}$	Roughness change likely due to ploughing
$\Delta VH_{HR} > th_{HR}$ and $-th_{MR} < \Delta VH_{MR} < th_{MR}$	Soil moisture change likely due to precipitation
$\Delta VH_{HR} < -th_{HR}$ and $-th_{MR} < \Delta VH_{MR} < th_{MR}$	Roughness change probably due to seedbed preparation after ploughing

1.3.3 Coherent change detection (CCD) for agricultural activities

The main disadvantages of the previous techniques rely in their own limitation of using only SAR backscatter (intensity or amplitude) and do not exploit information concerning the phase of the radar wave, which is a valuable piece of information (Robertson et al., 2023). Interferometric SAR (InSAR) coherence is a measure of the signal decorrelation between pairs of consecutive SAR acquisitions, making it a great metric for change detection (De Vroey et al., 2021).

Coherent change detection is a technique that uses both intensity and phase information to detect change between two SAR images. For example, an image acquired before tillage practices and an image acquired after tillage would be decorrelated in phase and intensity (Robertson et al., 2023). This method has demonstrated better sensitivity to small roughness changes (Preiss and Stacy, 2006) but requires that image pairs are acquired in the exact same geometry. Coherence (γ) can be computed as follows (Khabbazan et al., 2019; Voormansik et al., 2020):

$$\gamma = \frac{|\langle I_1 I_2^* \rangle|}{\sqrt{\langle I_1 I_1^* \rangle \langle I_2 I_2^* \rangle}} \quad (1.2)$$

where I_1 and I_2 are two complex SAR images (from different dates), $\langle \rangle$ indicates an ensemble averaging and $*$ is set as the complex conjugate. Values of coherence always range

from 0.0 (total decorrelation) to 1.0 (total correlation).

CCD is thus applied to two consecutive acquisitions and the output results of a change product with lower coherence values for tilled or harvested lands between image acquisitions (Robertson et al., 2023). Let's still keep in mind that harvesting and tillage are not the only field activities that can create decorrelation.

Several studies have proven that change in InSAR coherence levels, either a loss from a high level or a notable gain from a lower level, are closely related to substantial change in the surface structure of agricultural fields induced by land and soil management activities such as seedbed preparation during the spring season or harvest in the later fall season (Shang et al., 2020; Nasirzadehdizaji et al., 2021). A recent study which focused on the detection of mowing events on grassland parcels (De Vroey et al., 2021) developed three jump detection methods for coherence time series (Figure 1.4).

Mean shift (Figure 1.4a): This first method consists in a sliding averaging window of maximum size d . It allows to smooth the signal while still highlighting important shifts through time such as jumps. Every raw value T_i is replaced by the average of $[T_{i-r_1}, T_{i+r_2}]$ resulting in a smoothed value $ms(T_i)$, with r_1 and r_2 being respectively the number of points before and after, included in the window and adapted depending on the local variance of the signal. A jump is detected when the difference between $ms(T_i)$ and $ms(T_{i+1})$ exceeds a certain threshold.

Linear regression (Figure 1.4b): With this second method, time series are being smoothed by a linear regression on an asymmetrical sliding window. In this case, each raw value T_i is compared to the previous smoothed values $lr(T_i)$ obtained by the linear regression on $[T_{i-d}, \dots, T_{i-1}, T_i]$ allowing to detect sudden increases when compared to the preceding signal trend. Again, a jump is considered when the difference between the raw and the previous smoothed value exceeds a fixed threshold.

Two means (Figure 1.4c): This last jump detection method is based on a statistical hypothesis with a sliding window of even size d . The null hypothesis H_0 is the approximation by the mean value of the complete window $mean([T_1 : T_d])$. The alternative hypothesis H_1 is the approximation by a "step model" creating two means for the first $[T_1 : T_{d/2}]$ and second half $[T_{d/2+1} : T_d]$ of the window. A statistical function can be computed as follows:

$$F_{stat} = \frac{\left(\frac{SSR_0 - SSR_1}{k_1 - k_0} \right)}{\left(\frac{SSR_1}{d - k_1} \right)} \quad (1.3)$$

where $SSR[0, 1]$ and $k[0, 1]$ respectively are the Sum of Squared Residuals and the number

of parameters of $H_{[0,1]}$. A probability density function is then computed and a p -value is retrieved for each date. A p -value below a certain threshold results in the rejection of the null hypothesis H_0 . In that case, and if the following mean value is higher than the the first, a jump is detected.

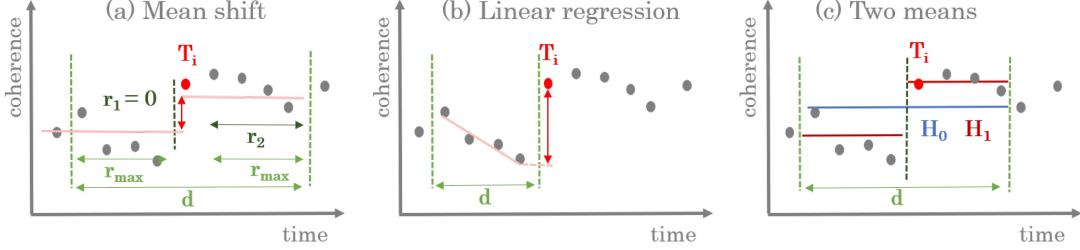


Figure 1.4: Visual representation of the three coherence jump detection methods where a jump occurs at T_i . **(a)** The mean shift method with a window size (d) of 9, where the difference between the smoothed values is evaluated by the red arrow, **(b)** the linear regression method of window size (d) of 4, where the raw value T_i is compared to the preceded smoothed value and **(c)** the two means method with a window size (d) of 8 comparing the null (H_0) and alternative hypothesis (H_1) (De Vroey et al., 2021).

Even though all three methods showed similar performances and robustness, mowing detections based on VV coherence smoothed by linear regression with a window of $d = 6$ were slightly more precise and accurate.

Similarly to these three jump detection methods, coherence downfalls can also be detected using the same principle.

1.4 Algorithms for crop classification

During past decades, a large amount of methods for classification have been used for scientific research purposes. Among them, K-Nearest Neighbour (KNN) (Cover and Hart, 1967), Suport Vector Machines (SVMs) (Cortes and Vapnik, 1995) and classification trees ensembles such as Random Forest (RF) (Breiman, 2001) are some of the most used techniques.

1.4.1 K-Nearest Neighbours: simple yet powerful

The k-Nearest Neighbour algorithm classifies new data points based on the majority class of their nearest neighbours in the feature space (Mucherino et al., 2009). The model, like many other classification algorithms, consist of two distinct phases: the training phase and the prediction phase. The training phase is quite different from other classification algorithms since it doesn't build an explicit model from the training dataset, it simply memorizes its features and class labels. The prediction phase consists of the assignment

of a certain class to an unlabeled data point. For this phase, the distance between the new point and all existing points in the dataset is computed by the algorithm. Equations 1.4, 1.5 and 1.6 highlight some of the distance metrics that can be used with different impacts on the final classification.

$$\textbf{Euclidean: } d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1.4)$$

$$\textbf{Manhattan: } d(x, y) = \sum_{i=1}^n |x_i - y_i| \quad (1.5)$$

$$\textbf{Minkowski: } \left(d(x, y) = \sum_{i=1}^n (|x_i - y_i|^p) \right)^{1/p} \quad (1.6)$$

with $d(x, y)$ the distance between the two points, x_i and y_i the coordinates of these two points, n the number of dimensions and p an integer parameter (Haviluddin et al., 2020).

The algorithm then identifies the k -nearest neighbours and assigns the class label of the majority of them to the new data point. Figure 1.5 shows an example of the prediction phase class assignment, where B is the new unknown point.

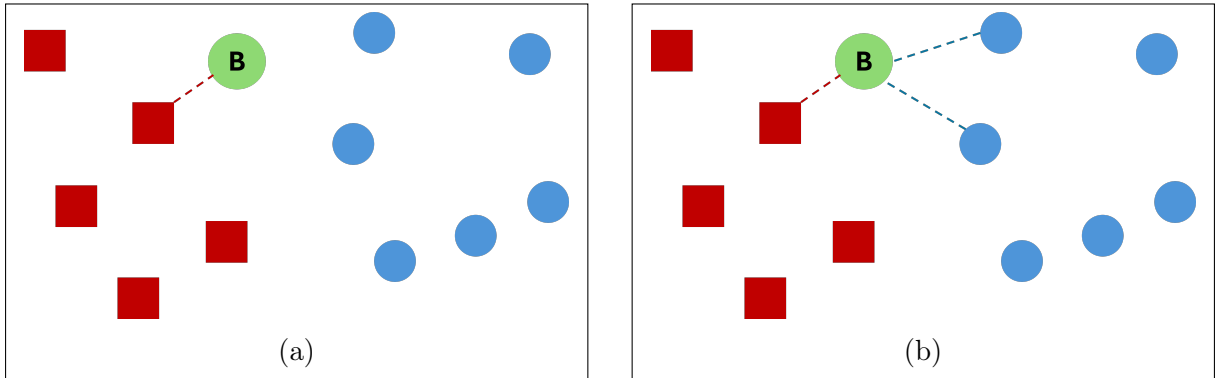


Figure 1.5: Schematic 2D representation of the KNN classification process with two classes. **(a)** A 1-NN decision rule where the new data point B is assigned to the red square class and **(b)** a k -NN decision rule where $k = 3$ and B is assigned to the blue circles class.

1.4.2 Support Vector Machine: high-dimensional spaces

The SVM classification process works by finding the optimal hyperplane after projection in a high dimensional space that best separates different classes in the feature space (Zhang, 2012). As for KNN classification, the process starts with collecting labeled training data from which a feature selection process will be performed to keep only features relevant for distinguishing different classes. The training phase consists in a first

part where selected features are mapped into a higher-dimensional space using a kernel function, the second part then aims to find the hyperplane that maximizes the margin between the classes. This hyperplane is the decision boundary that best separates the classes. Support vectors are the data points that lie closest to the hyperplane and are crucial in defining the optimal hyperplane. The decision function of the SVM then calculates the class label of a new data point based on its position relative to the optimal hyperplane. Figure 1.6 allows a clear and visual representation of the mentioned characteristics of a SVM classification.

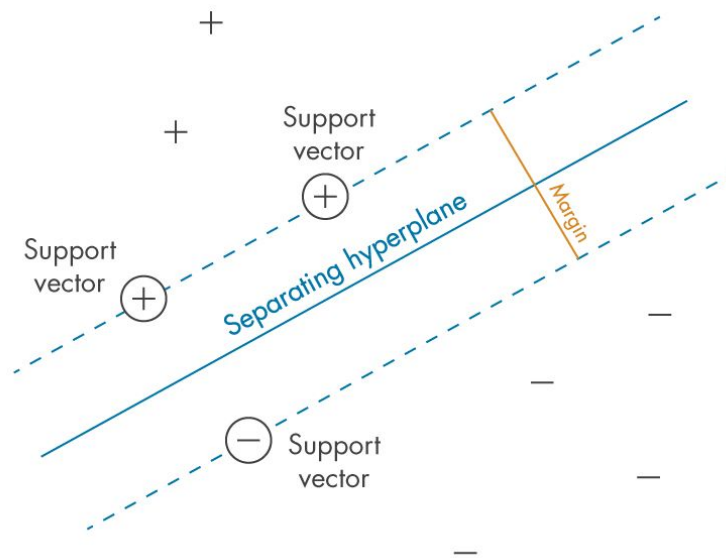


Figure 1.6: Schematic visualisation of the main SVM characteristics (Mathworks, n.d.).

The SVM algorithm handles non-linear classification tasks by using different kernel functions like polynomial, radial basis function (RBF), or sigmoid to map data into higher-dimensional spaces where linear separation is possible. Although the SVM classification method is particularly designed for binary classification, it can be extended using techniques like one-vs-one or one-vs-all to handle multiple classes efficiently.

1.4.3 Random Forest: a decision tree ensemble

Random Forest is a powerful ensemble learning technique which constructs numerous decision trees by leveraging random subsets of the dataset and features (Biau and Scornet, 2016). Firstly, the algorithm randomly selects samples from the dataset with replacement. This process is called bootstrap sampling or bagging (short for bootstrap-aggregating). Each decision tree of the model is then trained on a different bootstrap sample. For each one of the trees, at each node, a random subset of features is selected from the total feature set in order to introduce diversity in the trees and prevent overfitting. Each decision tree is then grown and stops when its maximum depth is reached, involving recursive data

splitting based on the most informative features. During the prediction phase, each tree classifies the input data, and the final prediction is determined by majority voting among all the trees in the forest. Figure 1.7 gives a schematic representation of the global functioning of a random forest classifier.

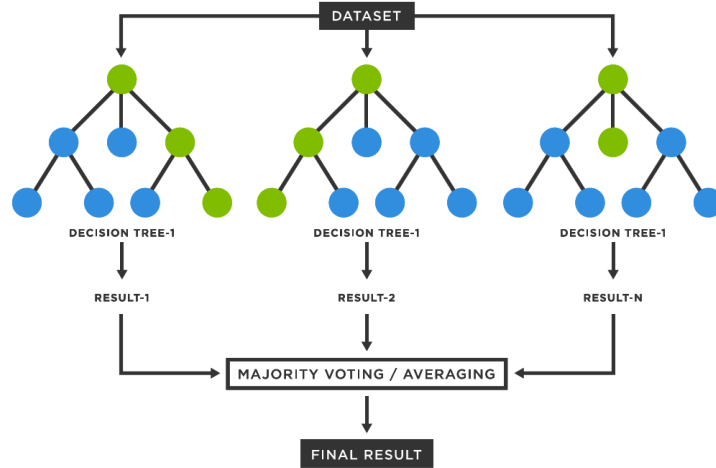


Figure 1.7: Schematic visualisation of the process of a Random Forest algorithm classification (Gunay, 2023).

1.4.4 Comparison between KNN, SVM and RF: pros and cons

Table 1.4 compares the three above-mentioned classification methods, highlighting their respective advantages and disadvantages (Boateng et al., 2020).

Table 1.4: Advantages and disadvantages of K-Nearest Neighbour, Support Vector Machine and Random Forest classification algorithms (Boateng et al., 2020).

Algorithm	Advantages	Disadvantages
K-Nearest Neighbors	• Simple to understand and implement • No training phase • Effective for small, well-structured datasets	• Slow inference time, especially with large datasets • Sensitivity to irrelevant features and noisy data • Difficulty in setting optimal K value
Support Vector Machines	• Effective in high-dimensional spaces • Robust to overfitting and noise	• More complex hyperparameter tuning • Less interpretable compared to simpler models • Can be computationally intensive with large datasets
Random Forest	• Robust to noise and overfitting • Good performance with small datasets • Less sensitive to small changes in training data • Scalable for large datasets	• Can be sensitive to hyperparameter tuning • Not as interpretable as some other models • Slower inference time compared to simpler models

1.4.5 Accuracy metrics for classification

When computing classifications, different accuracy metrics can be extracted to evaluate the classifier’s performance (Harikrishnan, 2019). These metrics are computed depending on the amount of True Positive (TP), True Negatives (TN), False Positives (FP) and False Negatives (FN). Among the possible accuracy metrics, Overall Accuracy (OA) stands out as the most common one and is computed following Equation 1.7.

$$\text{Overall Accuracy} = \frac{TN + TP}{TN + TP + FN + FP} \quad (1.7)$$

However, overall accuracy isn’t a perfect metric, especially when handling unbalanced

data. Overall accuracy does not differentiate between the types of errors (false positives and false negatives) and treats all misclassifications equally.

To provide a more detailed analysis of the classification’s performance, precision (Equation 1.8) and recall (Equation 1.9) can be computed. Precision measures the positive predictive value, approaching 1 as the number of false positives decreases. Recall (or sensitivity) indicates the true positive rate, approaching 1 as the number of false negatives decreases.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1.8)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (1.9)$$

Ideally, in a good classifier, both precision and recall need to be close to one, which also means that FP (false positives) and FN (false negatives) are near zero.

F1-score (Equation 1.10) takes into account precision and recall by computing a harmonic mean and is high only when both of the metrics are high. This makes F1 a great metric for accurate classification.

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (1.10)$$

Even though above described metrics are based on binary classification, the same principle can be easily extended to multi-class classification using the one-vs-rest approach to calculate per class metrics.

1.5 Conclusion on current scientific knowledge

This literature review highlights the pressing agricultural challenges of balancing sustainable practices with maintaining high yields to meet global food demands. The evolution and advancements in remote sensing technologies, particularly SAR, offer promising solutions for monitoring soil preparation activities and crop classification. Various change detection methods, notably InSAR coherence, have been extensively employed to track temporal changes in soil conditions. Moreover, the diversity of classification algorithms, each with its own advantages and disadvantages, offer useful insights into their applicability in early-season crop classification through SAR-derived soil properties data.

OBJECTIVES

The main objective of this master thesis is to assess the individual and combined contributions of X-band and C-band SAR data for seedbed preparation detection and early-season crop classification for the most common spring crops growing in the Belgian southern region.

Being able to retrieve such information in the early stages of the season can be very useful for multiple reasons. Firstly, being able to detect sowing dates could assist farm advisory services in monitoring and quantifying sowed areas in the early stages of the year, particularly during years when all sowing schedules are disrupted by weather events. The 2024 Belgian Crop Growth Monitoring (BCGMS) early-season report (BCGMS, 2024) emphasizes that spring cereal, maize, beet and potato sowing activities have been significantly delayed due to intense rain episodes in late April and May. Secondly, some agricultural models (e.g. decision support tools) could use this information to promptly assist farmers. This could be the case for the model predicting the yield, the nitrogen uptake by the plant, or disease vulnerability. In addition to improve food security by allowing better crop modelling, the goal is also to gain independence from administrative farmers' declarations, which are often released too late in the season to provide guidance on soil management practices. Moreover, multiple stakeholders related to the agricultural sector can find a lot of interest in such information. For example, agricultural insurance companies could use the detected sowing period to verify if farmers did or not follow the classic agricultural calendar in case of yield losses. This can also be interesting on a political aspect, mainly as part of the CAP of the EU, to monitor conservation agriculture practices and potentially improve or build new agricultural policies.

The following chapters will mainly be structured around the two main objectives of detecting seedbed preparation and classifying crops early in the growing season.

MATERIALS AND METHODS

This chapter presents a detailed explanation of materials and methods used to achieve the previous objectives. It starts by giving context to the research area, describes the data, and walks through the steps taken to process it, ultimately leading to the computation of the final results.

3.1 Study area: Beglian's silty region

The study is performed on agricultural fields in Wallonia, Belgium's Southern region covering 16 901 km² of the country. In 2022, Used Agricultural Area (UAA) covered 738 927 ha in Wallonia which represents 44% of the region and 54% of the country's total UAA (SPW Agriculture, n.d.). More specifically, focus is set on agricultural fields located in the area of Gembloux, Walhain and Perwez (Figure 3.1) mostly for proximity reasons for in-situ data collection (see Section 3.2.1). Additionally, the agricultural silty region where these cities are located is also a very dense region in terms of agricultural fields as seen in Section 1.1.1, Figure 1.1. Since the study focuses on the use of both C-band and X-band SAR imagery, the study area will be set at the intersection of both TerraSAR-X and Sentinel-1 (sensors used for the study, see Section 3.2) footprints (Figure 3.1), resulting in an area of 450 km².

The silty region is the largest agricultural region in Wallonia, covering 3 941 km². This region has very rich and fertile soil which is the main reason of its agricultural density compared to other regions of Wallonia (Figure 3.2). This area is known for its intensive cultivation of major crops like cereals, sugar beet and potatoes with half of the cereal production in Wallonia taking place in the silty region. In 2022, the silty region counted 4 652 farms, of which 3 920 are classified as professional, representing 84%. Among these, just over half specialize in large-scale crop farming, while 18% combine large-scale crop farming with livestock farming (SPW Agriculture, n.d.).

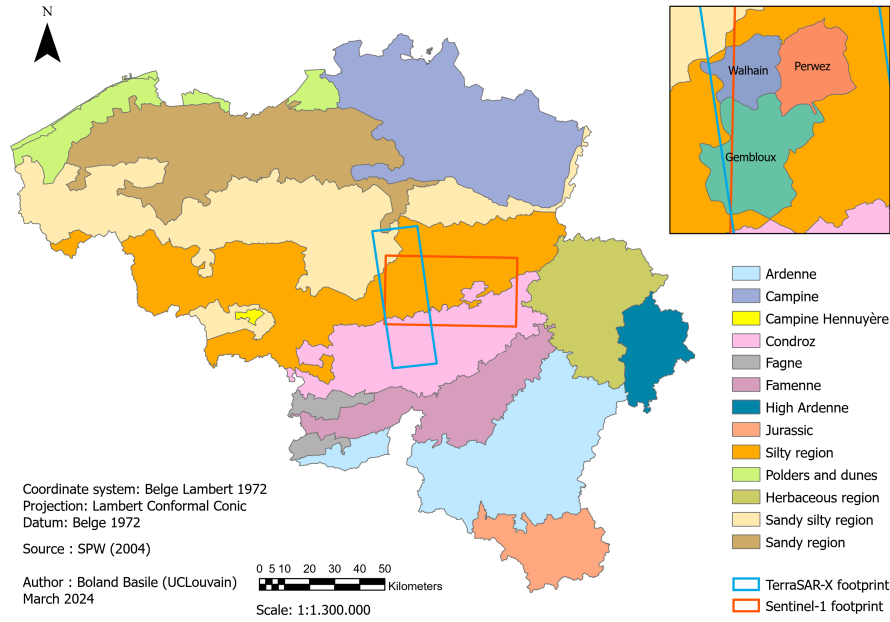


Figure 3.1: Map of agricultural regions of Belgium with location of the area of interest (AOI) in the Walloon region near the cities of Gembloux, Walhain and Perwez at the intersection of TerraSAR-X and Sentinel-1 footprints in this area. The AOI mainly covers the silty region but also presents small bits in the Condroz and the sandy silty region.

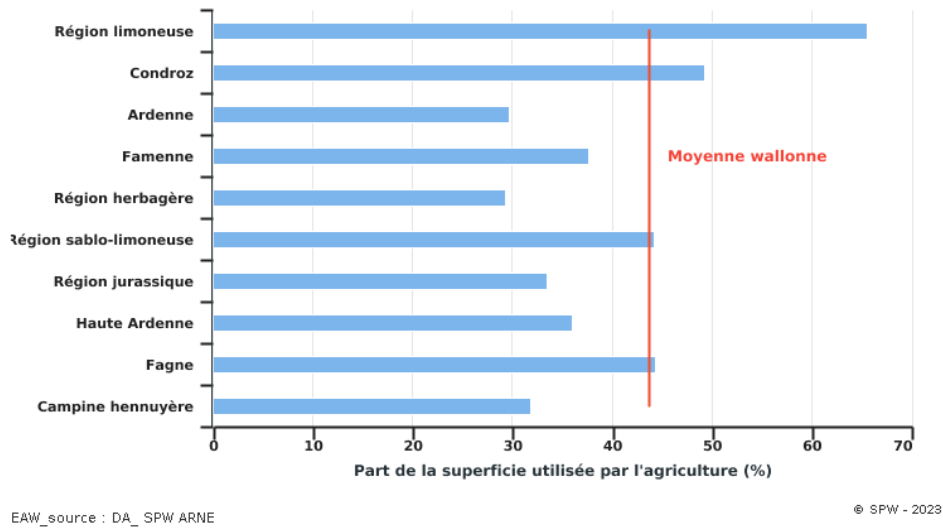


Figure 3.2: Share of total area used for agriculture by agricultural regions in 2022. The silty region is represented at the top of the chart and the vertical red line is showing the mean value for the Walloon region (SPW Agriculture, n.d.).

3.1.1 Soil characteristics: fertility in the silty region

The main soil types composing the area of interest are either silty soils with favorable natural drainage or silty soils with moderate or imperfect natural drainage (SPW, n.d.). Silty soils are very fertile and offer several benefits for agriculture. Thanks to their medium-sized particles, silty soils provide better water retention than sandy soils which al-

lows them to retain moisture for longer periods of time during drought. Moderate drainage also allows preventing waterlogging and promotes healthy root development (Clark, 2023).

On the other hand, silty soils can also be easily compacted due to their fine particles which reinforces the need to reduce tillage activities. Fortunately, addition of organic matter can help improve structure and fertility resulting in lower compaction and improving soil aeration (Warren and Taylor, 2017).

3.1.2 2021 agricultural season: a cold spring and wet summer

Since the study primarily centers on the spring period of 2021, during which in-situ data was gathered (refer to Section 3.2.1), climate details are more specifically considered for that specific timeframe.

The 2021 spring season has been a very cold season with an average temperature of 8.8°C compared to a normal season average of 10.5°C between 1991 and 2020 (Figure 3.3). With temperatures going from -2.7°C (06/03) to 25.6°C (09/05) in Uccle, the spring period of 2021 has been the third coldest spring since 1991 (IRM, 2021).

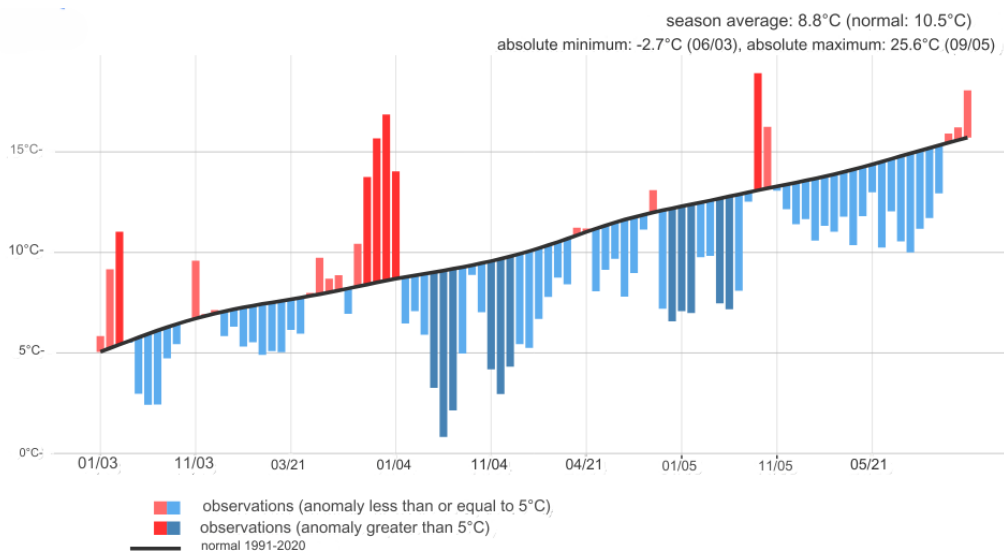
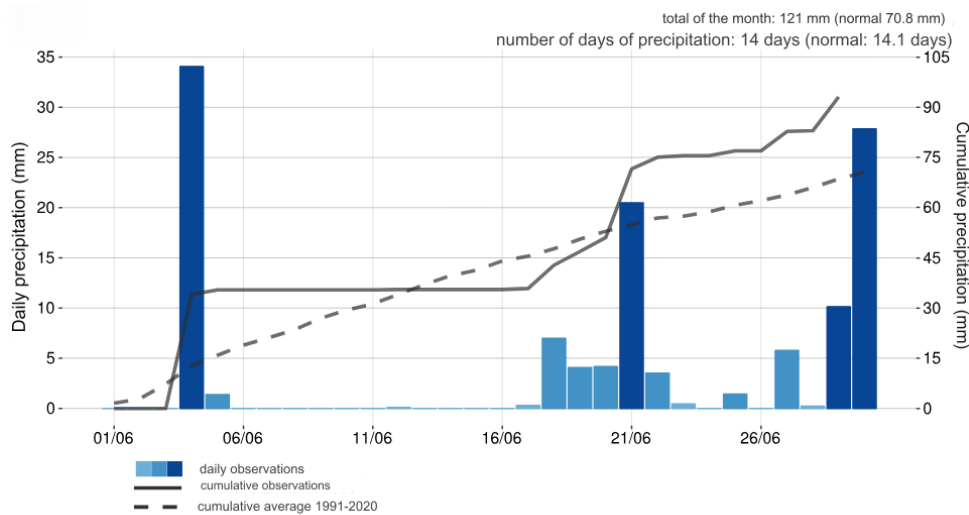


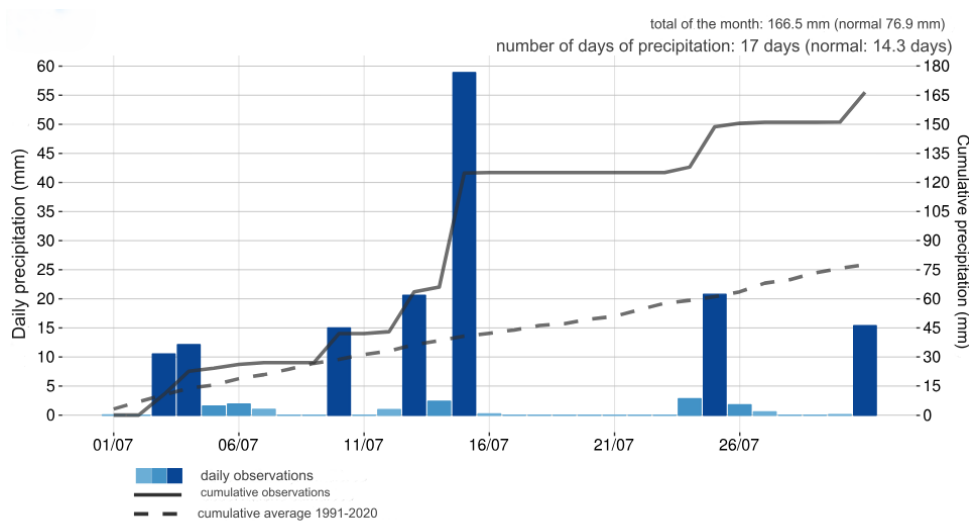
Figure 3.3: Average daily temperature in Uccle for the 2021 spring season. The black line indicates the normal average temperature between 1991 and 2020. Red and blue bars respectively indicate higher or lower anomalies (IRM, 2021).

The overall rainfall during the spring of 2021 pretty much matched the usual cumulative average of 165.6 mm in Uccle, but there were some notable monthly differences (Appendix A Figure 8.1). March and April turned out drier than normal, while May brought a lot more rain (IRM, 2021). On the other hand, June and July have been extremely wet months with some intense rain periods (Figure 3.4). These abundant rains triggered numerous flooding disasters not only in Belgium but also in neighboring Western

European countries such as France, the Netherlands, and western Germany. The total flooded area in Western Europe was estimated to be around 1.92 km², with agricultural land accounting for 64.4% of it (He et al., 2022). These events had significant short and long-term repercussions on agriculture. Grain, rapeseed, and flax crops were swept away, grain quality suffered, and harvests were unsuccessful due to flooding (He et al., 2022).



(a) June 2021



(b) July 2021

Figure 3.4: Daily and cumulative precipitation in Uccle for June 2021 (a) and July 2021 (b). Blue bars are representing daily observations with darkness depending on precipitation intensity. The dashed line shows the cumulative average between 1991 and 2020 while the full line shows cumulative observations (IRM, 2021).

As for sun exposure, the spring season of 2021 was slightly sunnier than usual, totaling 535 hours and 45 minutes of sunstroke compared to the typical average of 495 hours and 19 minutes in Uccle (IRM, 2021).

3.2 Available data

3.2.1 In-situ data collection

An intensive field campaign was conducted in 2021 by two researchers from the Geomatics Research team of UCLouvain, Diane Heymans and Jean Bouchat, to collect reference data on 170 fields between March 26th and May 31st (Figure 3.5). Seven field visits were conducted during this period, approximately 11 days apart, to align with the revisit period of TerraSAR-X. During each visit along the pre-established itinerary, fields were observed from the road, and the cover status of each parcel was recorded as either "covered" or "bare soil". In case of "bare soil" status, three possibilities could occur: soil was in a management phase (either "deep ploughing" or "prepared seedbed" state), presence of crop residuals or visibility of mounds. In that last case, details regarding shape, height, and line spacing of the mounds were collected. Finally, as soon as crop type was distinguishable, it was noted with four "principal" crop types being beets, potatoes, maize and wheat.

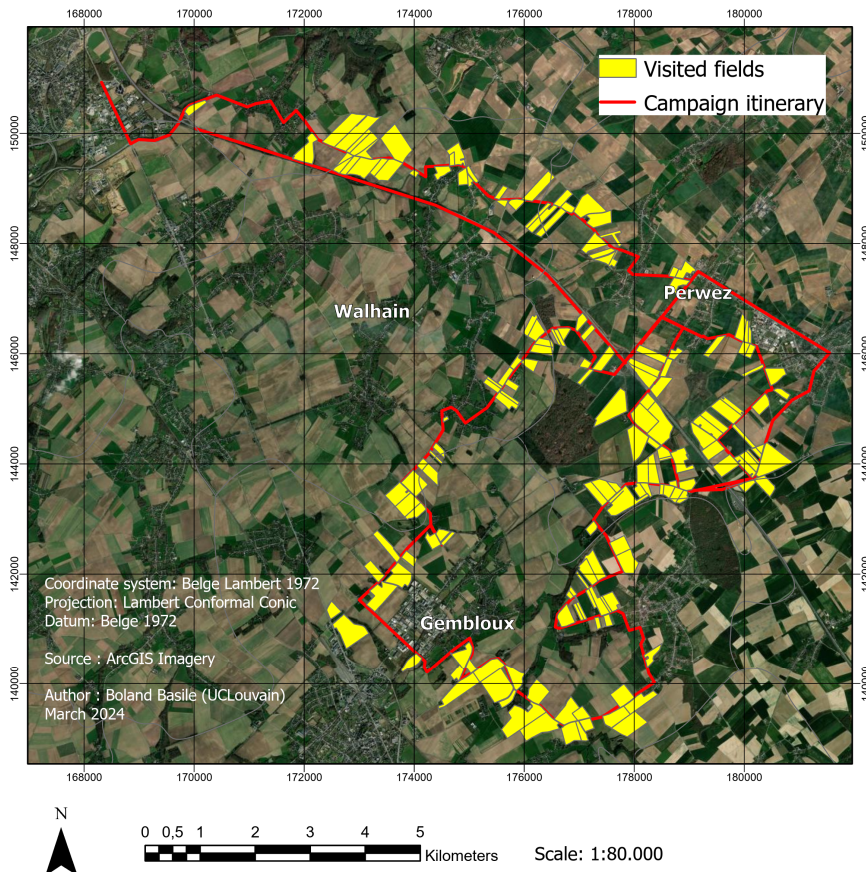


Figure 3.5: Map of visited fields (yellow) during the 2021 field campaign highlighting the campaign itinerary (red).

A Land Parcel Identification System (LPIS) is a digital system used to record all agri-

cultural parcels in the European Union Member States. It is a crucial tool in the CAP for verifying eligibility for area-based subsidies, which totaled around 45.5 billion euros in 2015 (European Court of Auditors, 2016). The LPIS dataset assigns an ID to each parcel and also provide useful information concerning the parcel use and cover type. The anonymized registry of agricultural parcels can vary from year to year and is accessible annually on the Wallonia geoportal (Geoportail Wallonie, n.d.).

Since the field campaign was conducted in spring 2021, only the 2020 LPIS data was available at that time of the year and fields were therefore identified following the previous year dataset. As parcels not only change IDs from year to year, their shape can vary resulting in parcels being smaller, larger, split or merged from a year to another. Transition from 2020 LPIS reference data used for the field campaign to 2021 LPIS data is made using a GIS software. By extracting the centroid of each 2020 parcel and spatially joining it to the 2021 LPIS layer, link could be made between the two datasets. This method worked well for small fields size changes and merged fields but had to be manually verified for fields being split in two or more fields from 2020 to 2021. In that case, the 2021 field presenting the same cover as the 2020 field was kept. If split fields were covered by the same culture in 2021, either the closest field to the road or largest field was kept. After presented processing steps, the final field campaign dataset consists of 166 parcels allowing a clear link between 2020 and 2021 parcels identification numbers.

From the 166 visited fields, 38 are classified as beets, 26 as potatoes, 23 as maize and one as spring cereal (oat) in the Walloon LPIS. The 78 remaining fields are either winter crops or less common cultures.

3.2.2 SAR imagery: X-band and C-band

As previously mentioned, the aim of this study is to evaluate the contribution of C-band and X-band SAR imagery for seedbed preparation detection and early-season crop classification. In this case, chosen sensors are TerraSAR-X for X-band imagery and the Sentinel-1 constellation for C-band imagery.

TerraSAR-X

TerraSAR-X is a radar satellite born from the partnership between DLR (the German Aerospace Center) and Airbus Space and Defense, launched in June 2007 and who began its operations in January 2008. With a wavelength of 31 mm (frequency of 9.65 GHz), TerraSAR-X is an X-band SAR sensor capable to record data with a spatial resolution up to 1 m, regardless of weather and lighting conditions (ESA, n.d.b). The satellite follows a sun-synchronous dusk-dawn orbit at an altitude of 514 km allowing a revisit period of 11 days.

A total of 10 acquisitions, were specifically ordered and acquired from March 26 to the third of July 2021 in Stripmap Mode (SM) with a Ground Sampling Distance (GSD) of 3.5 m and an incidence angle ranging from 49.03 to 49.90 degrees. The signal is recorded in dual polarization mode, covering VV and VH polarizations. The swath width of the products is 30 km for a length of 50 km.

Products are pre-processed at EEC level (Enhanced Ellipsoid Corrected) which corresponds to the highest level of processing available. Reference points and a Digital Elevation Model (DEM) are applied for the geocoding allowing a quick and highly accurate overlay with other spatial data (Airbus, n.d.). In addition, a Radiometric Enhancement (RE) is applied, which intentionally decreases spatial resolution to reduce speckle (caused by the interference of the radar waves scattered back from the target area) and obtain a better radiometric resolution of about 1.5 dB (DLR, n.d.). Finally, radiometric calibration is applied using ESA’s Sentinel Application Platform (SNAP) (ESA, 2022) and the Sentinel-1 Toolbox (S1TBX) in order to convert the image pixel values from digital numbers (DN) to the standard geophysical measurement unit of radar backscatter (NASA, n.d.). Final products consist of two-band raster files respectively storing sigma nought (σ^0) values for VH and VV polarizations. Sigma nought stands as the normalized backscattering coefficient, which measures the intensity of the radar signal reflected back to the sensor.

Sentinel-1

The ESA Sentinel-1 constellation (Sentinel-1A & Sentinel-1B) from the EU Copernicus program operates in C-band at a centre frequency of 5.405 GHz (wavelength of 5.55 cm) (ESA, n.d.a). Sentinel-1 allows a revisit period of 6 days thanks to the twin sensors (S1A & S1B), which were still operational in 2021 (S1B’s mission has ended since December 23, 2021). Level-1 Single-Look Complex (SLC) acquisitions were collected between March 17th and July 15th in Interferometric Wide (IW) and descending mode and pre-processed using previous mentioned SNAP software. Products have a GSD of 10 m and signal is recorded in dual polarization mode (VV and VH). A total of 20 acquisitions are selected following the relative orbit number 37. Sentinel-1 data also offers valuable information on the phase of the returned signal. Allowing to compute interferometric coherence, a change metric presented in Section 3.3.2.

The pre-processing steps to obtain backscattering coefficients (σ^0) for VV and VH polarizations are applied as follows:

1. TOPSAR split;
2. orbit file application;

3. thermal noise removal;
4. calibration to σ^0 ;
5. TOPSAR deburst; and
6. TOPSAR merge.

3.2.3 Precipitation data

As soil moisture strongly influences radar backscattering (Hadria et al., 2009; McNairn et al., 1996) it is crucial to consider its precise monitoring in the study area. To achieve this, precipitation data is retrieved from the Agromet platform. Agromet was established by the Walloon Agronomic Research Center (CRA-W) in collaboration with the Royal Meteorological Institute (RMI) and universities of Gembloux and Louvain. The purpose of this online platform is to assist farmers in gathering real-time agro-meteorological data (precipitation, sub-soil temperature, wind speed and direction, etc.) at high spatial (1 km²) and temporal (hourly and daily) resolution (CRA-W, n.d.). Over 30 meteorological stations belonging to the PAMESEB network cover the Walloon region and historical data as well as predictions can easily be displayed and downloaded using the Agromet API or by directly downloading data from the website in Excel, CSV or JSON formats.

Based on the study area and period, precipitation data from the station located in Louvain-la-Neuve is collected, as it is the closest to the area of interest (~ 5 km), for the period of time starting in early March and ending in late July 2021.

3.3 Data processing

3.3.1 Field polygons

In order to reduce the influence of edge effects that could occur due to variations in terrain near the boundaries of each field, an inner buffer of 20 m is applied to each field polygon. By excluding these potentially noisy areas, the future analysis will focus on the core characteristics of each parcel, leading to more accurate and representative statistics. Additionally, applying this inner buffer helps to reduce possible pixel contribution from the edges of the polygon, ensuring that only complete pixels within the polygon contribute to the computation of statistics.

3.3.2 Zonal statistics, indices and coherence

The main extracted feature for each radar imagery pixel is sigma nought (σ^0), the normalized backscattering coefficient, which measures the strength of the radar signal reflected back to the sensor. The signal intensity is recorded for VH and VV polarization.

This section will detail the process to obtain the different features that will characterize each crop field through time.

Zonal statistics

Based on the above mentioned sigma nought backscatter coefficient, zonal statistics are performed on raster data derived from both TerraSAR-X and Sentinel-1 sensors (see Section 3.2.2) within the area of interest. Zonal statistics are computed through parcel averaging, extracting a single value for each parcel and enabling speckle reduction for the analysis. Using Python coding language, a function is created to compute summary statistics, including mean, standard deviation, median, minimum and maximum values for both VV and VH polarization by overlaying each radar imagery onto the agricultural field boundaries. These zonal statistics serve as a robust tool for extracting meaningful information from radar imagery and will be valuable for temporal evolution analysis. Let's note that backscattering coefficients are converted in dB, the common backscattering coefficient unit, following Equation 3.1.

$$\sigma^0 [dB] = 10 \times \log_{10}(\sigma^0) \quad (3.1)$$

In addition to computing summary statistics for each polygon, a criterion is implemented to reject polygons that do not include a minimum of 50 pixels of the radar image. This criterion ensures that the analysis focuses only on polygons with sufficient spatial coverage of the radar data. By excluding polygons that fail to meet this requirement, the analysis maintains a robust and meaningful representation of the parcels, enhancing the reliability of the derived statistics and insights. More discussion and comparison about this rejection criterion is presented in Section 5.2.1. Moreover, in order to capture the true temporal evolution of each field and as the sensors footprints can slightly move from one acquisition to another, it is important to exclude fields that aren't covered by every acquisition.

Radar indices

In addition to backscattering statistics, several radar indices can be computed based on the VV and VH mean backscattering coefficients. Among these radar indices, the Cross-polarization Ratio (CR) and Dual-polarization Radar Vegetation Index (DpRVI) are known to be very useful SAR vegetation dynamics indices (Vreugdenhil et al., 2020; Mandal et al., 2020). Equations 3.2 and 3.3 respectively indicate how CR and DpRVI are computed, with σ_{VH}^0 and σ_{VV}^0 already converted in dB.

$$CR [dB] = \frac{\sigma_{VH}^0}{\sigma_{VV}^0} \quad (3.2)$$

$$DpRVI [dB] = \frac{4 \times \sigma_{VH}^0}{\sigma_{VH}^0 + \sigma_{VV}^0} \quad (3.3)$$

Interferometric coherence

Coherence analysis is a fundamental aspect of this study, providing important insights into temporal changes occurring through time and space. Coherence is using intensity data of radar imagery and more specifically determined by the phase difference between two radar images acquired over the same area (Robertson et al., 2023). This metric serves as a strong indicator of the consistency of these phase differences, allowing the detection of subtle changes in surface features. The computation of interferometric coherence has been detailed in Section 1.3.3.

Similarly to the process applied to the backscattering coefficient, different statistics of coherence values are computed for each field including mean, standard deviation, minimum and maximum. This allows to extract spatial and temporal information about coherence and create more features for classification purposes. Despite its significance for change detection, coherence could only be applied to Sentinel-1 data for this study as the processing level (EEC) of the provided TerraSAR-X imagery didn't allow access to phase information, making coherence computation unfeasible.

3.3.3 Time series

Given the study strongly focuses on the early detection of seedbed preparation and crop classification, the temporal analysis of each field and crop type characteristic becomes of great importance. Time series analysis captures and allows to visualize temporal dynamics of a certain feature and compare it between different classes (here, crop types). This approach not only facilitates the detection of temporal trends or changes but also serves as a first insight on how data evolves through time, providing a better understanding of the temporal dynamic shaping the study area over time.

3.4 Seedbed preparation date detection

The first part of this work consist of the detection of the seedbed preparation period, closely related to the sowing date.

3.4.1 Coherence time series

Coherence is an easy to understand variable as it reflects change over time with a scale from 0 to 1 (0 being related to important change and 1 to totally unchanged surfaces). A sowing phase transitioning from a deeply ploughed to a prepared seedbed soil is subject to strong change in soil roughness and could be detected using coherence time series (Shang

et al., 2020; Nasirzadehdizaji et al., 2021). An important downfall in coherence should be detected transitioning from a deep worked soil to a smooth and more homogeneous parcel surface. This downfall can be detected by locating the coherence minima on the early-season period. As the coherence can only be computed when information on phase is available, focus will be set on parcels covered by the selected Sentinel-1 footprint, making only C-band data considered for this part of the study.

3.4.2 In-situ data

The field campaign data collected information on the management status for every visited parcel. This information, linked to the LPIS, allows to create some sort of soil management status evolution through time for each visit date during the campaign. Three different soil management status have been identified from March 26 up to May 31: "deep ploughing", "prepared seedbed" and "covered". In this case, the transition from "deep ploughing" to "prepared seedbed" is of interest meaning fields go from intensive soil work to seedbed preparation during that period of time of roughly 10 days (field revisit period during campaign). The "prepared seedbed" state for potato fields corresponds to the formation of potato hills.

By grouping fields by crop type, it is possible to define inside clusters depending on a common transition period. The visited fields estimated sowing period's distribution for each crop type is presented in Table 3.1. Unfortunately, no data was retrievable for spring cereals as no transition from intensive soil work to seedbed preparation was recorded for all of the visited fields. Moreover, considering that only one of the visited fields was categorized by the LPIS as a spring cereal field (oat, in this case), the representativeness would be very low regardless. Some fields could unfortunately not be assigned to a sowing period due to no clear recording of a transition from "deep ploughing" to "seedbed prepared" state.

Table 3.1: Number of fields being sown for each crop type (beet, potato and maize) between each visit during the 2021 field campaign. The period refers to the time between two field visits where soil management status transitions from "deep ploughing" to "prepared seedbed". The most dominant cluster for each crop type is highlighted with a different color. The date format is set as mm-dd.

Sowing period	Beet	Potato	Maize
03-26 to 04-07	23	0	0
04-08 to 04-19	1	0	2
04-20 to 04-27	1	12	6
04-28 to 05-08	0	7	1
05-09 to 05-20	0	1	0
Insufficient information/Outlier	13	6	11

3.5 Early season crop type classification

Next to seedbed preparation detection, early season crop type classification constitutes the second objective of this master thesis. The temporal aspect of this point will be studied with cumulative increase in the amount of input data given to the Random Forest model.

3.5.1 Feature extraction

The main purpose of the study is to have an accurate classification as early as possible in the season, again putting strong focus on the temporal aspect. This is why it is crucial to see how the classifier evolves depending on the amount of information given to the classification model. In other words, see how the classification results evolve through time after the first acquisition data is given.

To achieve this, it is necessary to sequence the features that will serve as input for the classification model. Indeed, as the aim is to evaluate temporal evolution each of the input features have to be gradually added depending on the acquisition date. The point is to run the classifier once with the first acquisition features, then again with the two first acquisition features, etc. until the whole range of acquisitions during the period of time of the study is covered. Input features for each acquisition are VV & VH σ^0 backscatter (mean and standard deviation), CR, DpRVI and VV & VH coherence for C-band data (mean and standard deviation) for a total of 6 and 10 features by parcel respectively from TerraSAR-X and Sentinel-1.

3.5.2 Random Forest classification

In the case of this study, the chosen classification algorithm is Random Forest. Random Forest stands out as an optimal choice for various reasons (Boateng et al., 2020). Firstly, its robustness to noise and overfitting makes it suitable for datasets characterized by noise or imbalance. Moreover, the algorithm exhibits notable performance even when dealing with small datasets and shows reduced sensitivity to minor fluctuations in the training data, a feature particularly advantageous compared to models like the KNN algorithm. Furthermore, it offers valuable insights into feature importance, facilitating model interpretation and fitting. Despite being slower relative to simpler models, Random Forest remains scalable for large datasets, thereby proving its applicability on wider areas.

The Random Forest algorithm is implemented in Python using the *RandomForestClassifier* class through the *Scikit Learn* library's *ensemble* module. The number of trees for each classification is set at 500 with a maximum depth of 20. The training and testing datasets are split following a 80/20 distribution.

Working with unbalanced data is a common challenge in real-life scenarios, where normal or common instances often vastly outnumber abnormal or uncommon ones (Chawla et al., 2002). The Synthetic Minority Over-sampling Technique (SMOTE) addresses this issue by generating synthetic examples of the minority class, effectively rebalancing the training dataset for the classification. Unlike traditional over-sampling methods, such as replication, SMOTE expands decision regions in feature space without risking overfitting (Chawla et al., 2002).

Finally, two important metrics are extracted for each classification: F1 score and the Mean Decrease in Impurity (MDI). While Overall Accuracy (OA) is very commonly used in scientific literature for classification, F1 score is often of application for imbalanced data (Huilgol, 2019). This metric is therefore preferable in this particular case. The MDI is a measure of variable importance in tree-based ensemble methods, such as Random Forests. It quantifies the contribution of each input variable to the overall predictive performance of the ensemble. Basically, a high MDI value indicates that a variable is more important for prediction, as it leads to greater reductions in impurity when used in splitting nodes during tree construction (Louppe et al., 2013).

RESULTS

The following chapter presents the results of the detailed study and is divided in three main parts: the time series analysis as data exploration, the seedbed preparation detection and the early-season crop classification performance analysis. While this chapter presents some minor discussion points for clarity, the majority of the results discussion, comparison and critical analysis will be addressed in Chapter 5.

4.1 Time series: a first visual analysis

Time series provide valuable visual aids for gaining a preliminary impression and understanding of the temporal behavior of a specific signal, variable, or feature over time. In this case, the focus is set on identifying potential differences in temporal evolution among different crop types, namely beet, maize, and potato in the area of interest.

4.1.1 TerraSAR-X time series

The dataset comprises 10 TerraSAR-X acquisitions and their derived features, including VV & VH backscatter, CR and DpRVI. Figure 4.1 illustrates the temporal behavior of the mean CR in decibels (left vertical axis) and the evolution of hourly precipitation in millimeters per square meter (right vertical axis) throughout the 2021 spring season. Other TerraSAR-X time series are presented in Figures 8.2 through 8.7 of Appendix B. The CR evolution reveals a globally similar trend for every crop type with noticeable differences at specific times of the season, particularly in early April when the CR for potato fields drops lower than for beet and maize fields, and at the beginning of June when all three crop types seem to diverge after experiencing a drop of approximately one decibel.

Precipitation data shows a significant peak, almost reaching 30 mm/m², in early June, with substantial precipitation also occurring throughout the rest of the month. The

correlation between CR values and precipitation is challenging to identify clearly due to the long revisit period of the TerraSAR-X satellite. Even though increase in CR can be observed after high precipitation events, the causal effect can't be fully confirmed only based on this information.

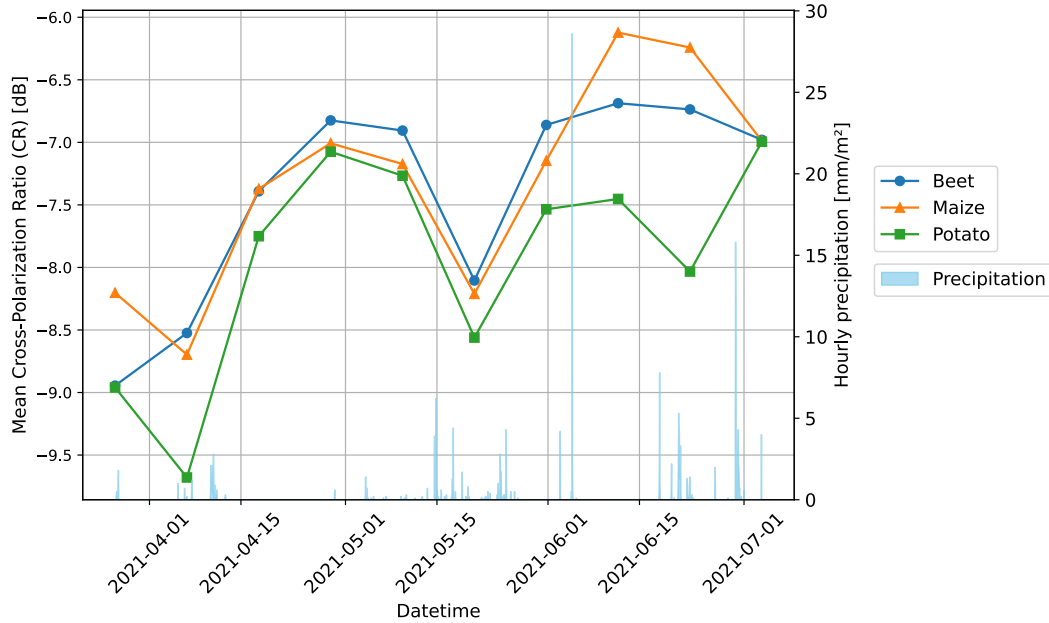


Figure 4.1: Mean evolution of the TerraSAR-X Cross-polarization Ratio (CR) and hourly precipitation through the 2021 spring season for all parcels of each studied culture (beet, maize, potato).

Plots showing the different TerraSAR-X mean time series with associated standard deviation are available in Appendix B (Figures 8.4 to 8.8).

4.1.2 Sentinel-1 time series

Figure 4.2 illustrates the mean evolution of the CR derived from Sentinel-1 backscatter information (left vertical axis) and hourly precipitation in mm per square meters (right vertical axis) through the 2021 season. Other time series features are available in Appendix B (Figures 8.9 to 8.13). The CR values for all crops exhibit noticeable increases and decreases at different times of the season, with an overall common trend in the early season characterized by a shared drop followed by a significant increase in early June. Precipitation data mirror those presented for TerraSAR-X time series, except that the Sentinel-1 time series extend further into the season, highlighting important precipitation events occurring in July. Whether the increase in CR values starting in June is related to precipitation or to vegetation development isn't clear to drop to specific conclusions.

Plots presenting the different Sentinel-1 mean time series with associated standard deviation are available in Appendix B (Figures 8.14 to 8.19).

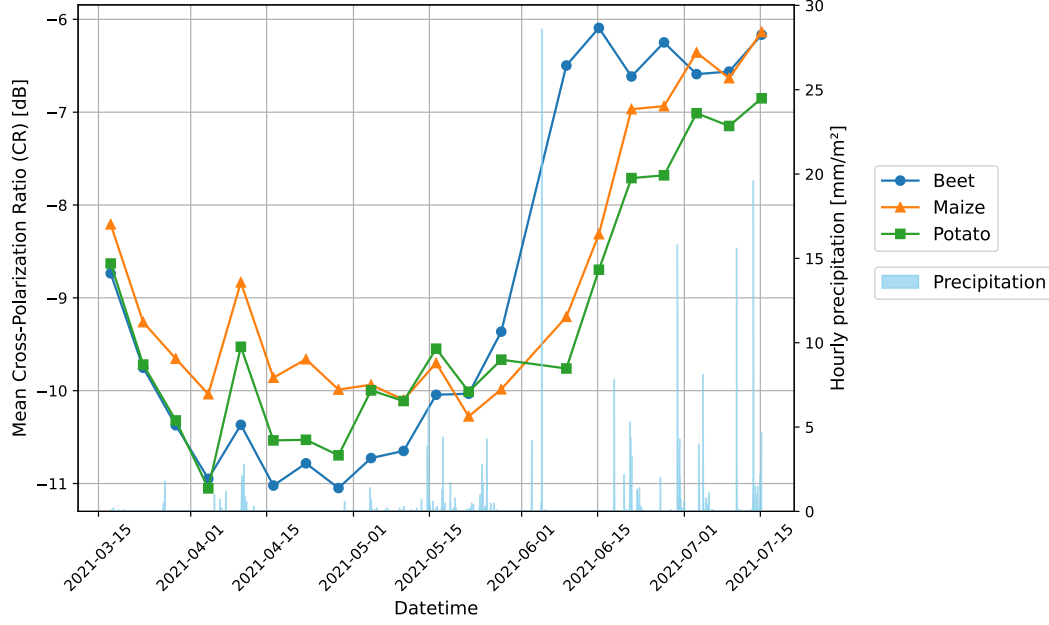


Figure 4.2: Mean evolution of the Sentinel-1 Cross-polarization Ratio (CR) and hourly precipitation through the 2021 spring season for all parcels of each studied culture (beet, maize, potato).

4.2 Seedbed preparation detection

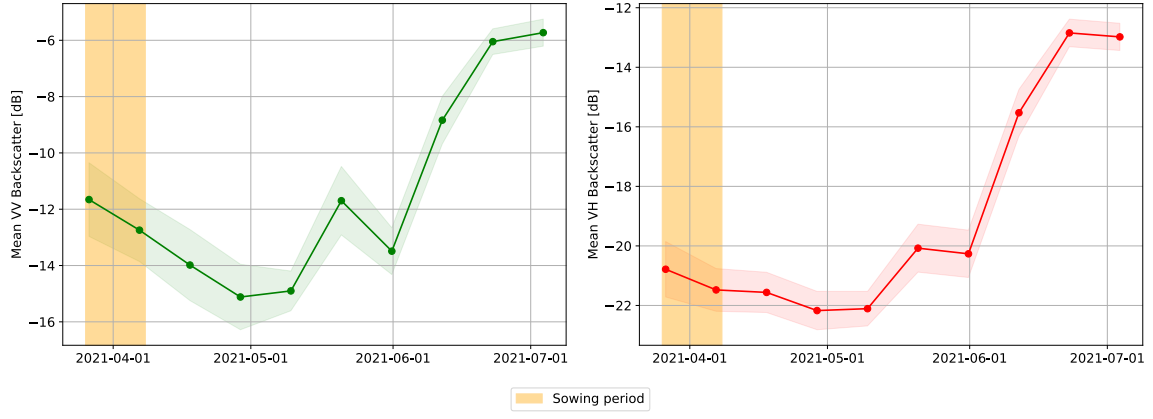
The seedbed preparation detection section will here present two main phases: a testing phase using field campaign data and time series from both sensors and a second phase that will extend results on a wider area.

4.2.1 Sample fields: TerraSAR-x & Sentinel-1

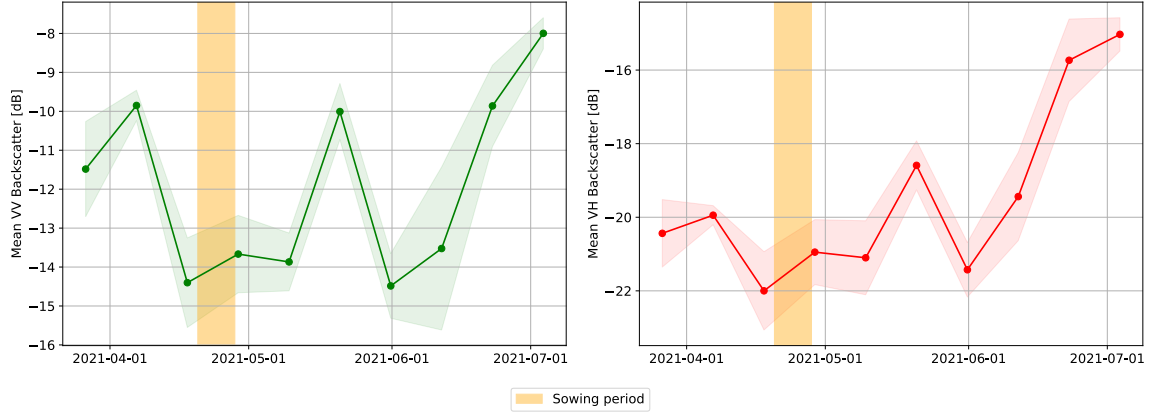
For this initial testing phase, only sample fields extracted from the field campaign are under consideration. Specifically, the focus is placed on visited fields that share the same sowing period according to the collected data. For each crop type, the most representative sub-group with regards to the sowing period has been retained, as previously highlighted in Table 3.1 of Section 3.4.2.

TerraSAR-X

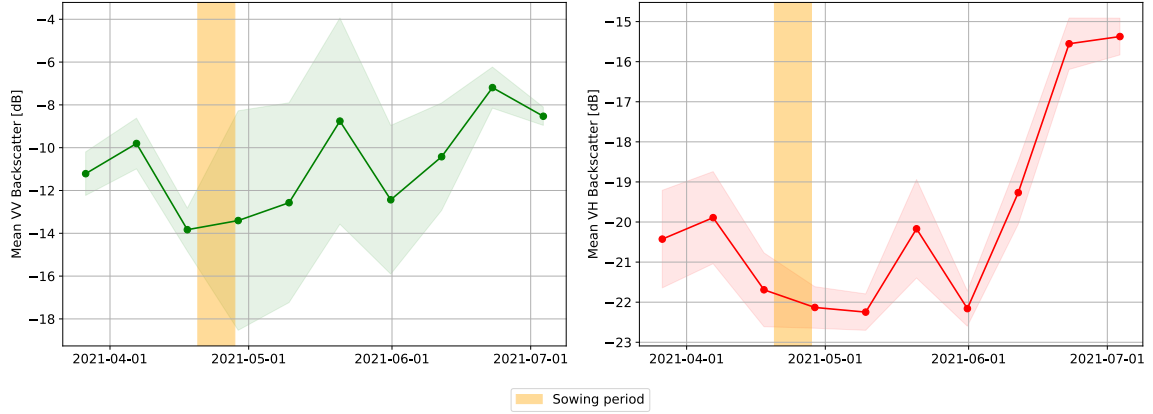
Figure 4.3 presents TerraSAR-X time series for mean VV & VH backscatter data in decibels respectively for beet, maize and potato fields of the field campaign. A buffer surrounding the line showcases the standard deviation around the mean value for each acquisition. The estimated sowing period extracted from the field visit campaign is showcased as an orange vertical stripe. The large revisit period between each acquisition does unfortunately not allow to create a clear link between that estimated sowing period and the signal evolution.



(a) Beet fields



(b) Maize fields



(c) Potato fields

Figure 4.3: Mean evolution and associated standard deviation of the TerraSAR-X VV (green) & VH (red) backscatter (σ_{VV}^0 & σ_{VH}^0) for all sample fields that share the same sowing period (vertical orange stripe) for each studied crop type.

Sentinel-1

Figures 4.4, 4.5, and 4.6 depict the same information, but applied to Sentinel-1 time series with the addition of VV & VH coherence, represented respectively in blue and purple. Backscatter values exhibit significant fluctuations over time translating the varying intensity of the returned signal, closely related with soil roughness and moisture.

These backscatter values present a local minimum occurring in the early season coinciding with the designated sowing period for each crop type, although beet fields show a lower backscatter value in late April. Conversely, coherence values tend to vary less significantly over time, still showing a noticeable drop during the sowing period, often followed by an increase. This trend is particularly pronounced in the VV coherence time series, where a substantial drop is immediately succeeded by a significant increase. This result isn't surprising knowing coherence is a metric of change between two acquisitions, an important downfall therefore highlights a disturbance in soil roughness and/or soil moisture, typically occurring when a crop field transitions from deep soil work to a prepared seedbed.

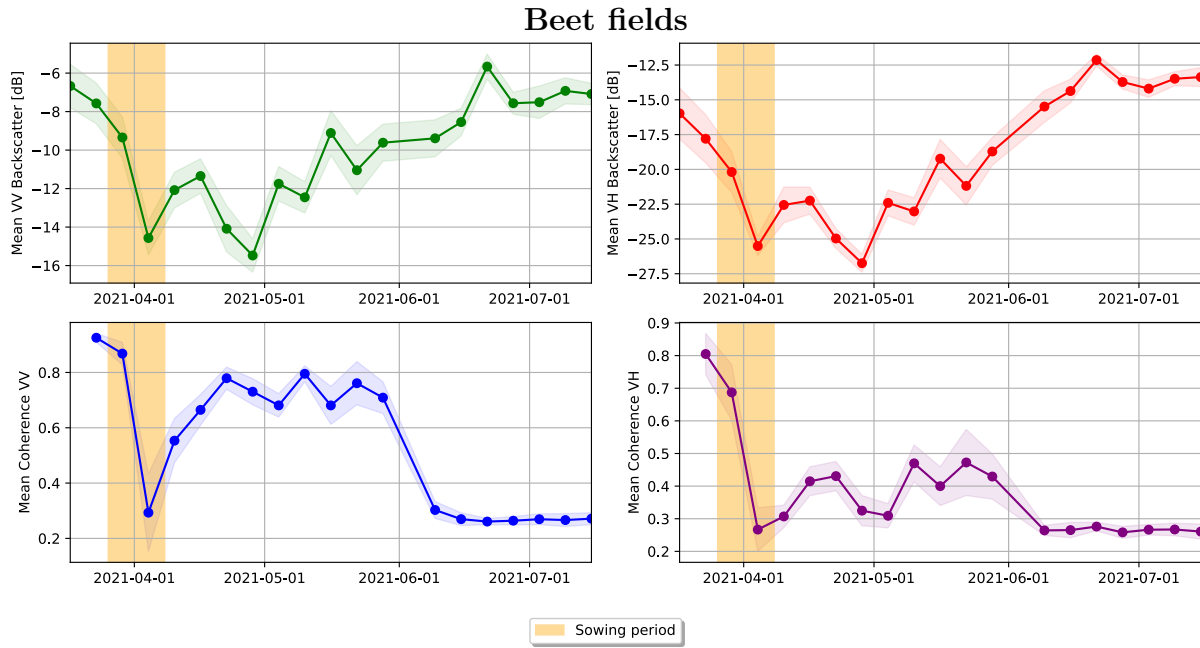


Figure 4.4: Mean evolution and associated standard deviation of the TerraSAR-X VV (green) & VH (red) backscatter (σ_{VV}^0 & σ_{VH}^0) and VV (blue) & VH (purple) coherence for all field campaign beet fields that share the same sowing period (vertical orange stripe).

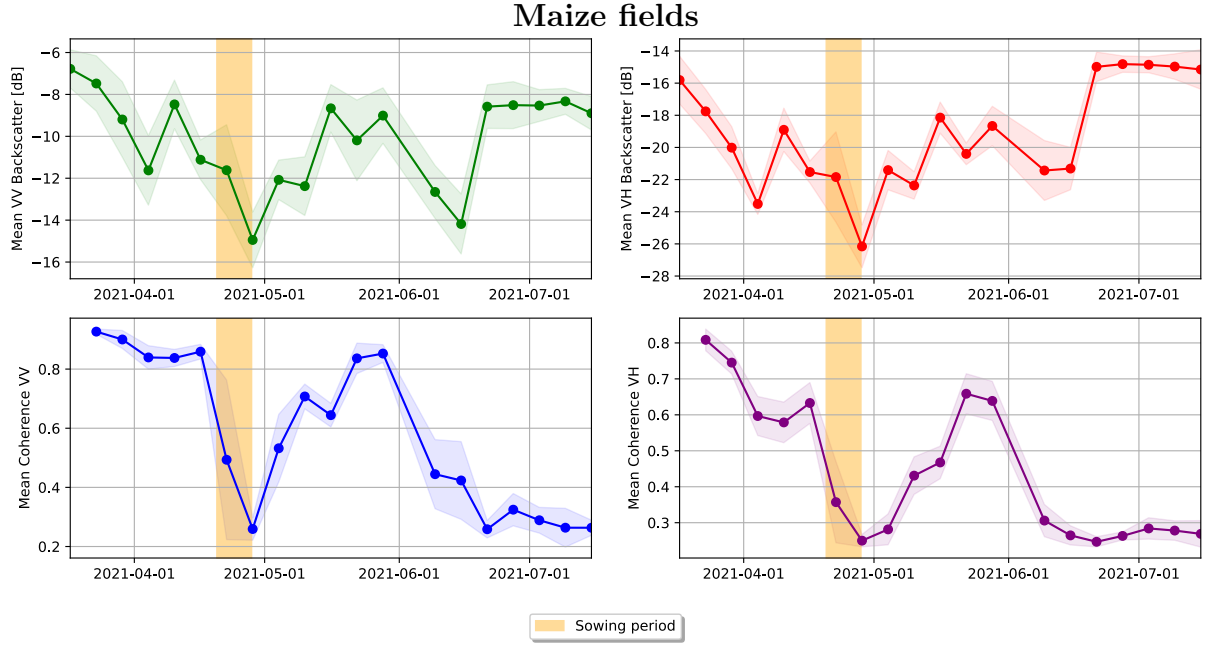


Figure 4.5: Mean evolution and associated standard deviation of the TerraSAR-X VV (green) & VH (red) backscatter (σ_{VV}^0 & σ_{VH}^0) and VV (blue) & VH (purple) coherence for all field campaign maize fields that share the same sowing period (vertical orange stripe).

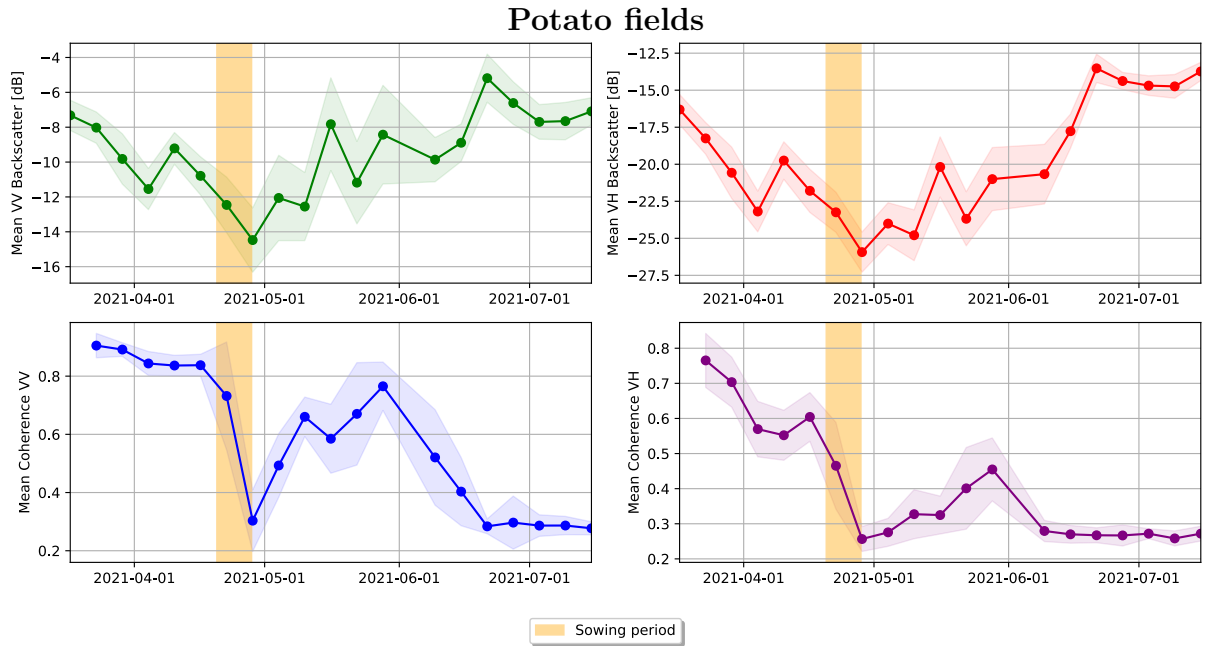


Figure 4.6: Mean evolution and associated standard deviation of the TerraSAR-X VV (green) & VH (red) backscatter (σ_{VV}^0 & σ_{VH}^0) and VV (blue) & VH (purple) coherence for all field campaign potato fields that share the same sowing period (vertical orange stripe).

4.2.2 Sentinel-1 footprint

Following the results presented in the previous section, the detection of seedbed preparation can be extended to the complete Sentinel-1 footprint selected for the study to visualize the spatial and temporal distribution of detected seedbed preparation dates for each crop type of interest. The applied principle involves extracting the date at which VV coherence was at its lowest point between March 15th and May 31st, considered as the interval of time in which sowing activities take place.

Figure 4.7 illustrates this spatial distribution on the sensor's footprint for beet fields, with color intensity indicating a later detected date in the season. The distribution plot in the lower part of the figure reveals a predominant number of fields being prepared for sowing in the first days of April. The distribution map exhibits a slight concentration of later detected fields in the South-West part of the footprint.

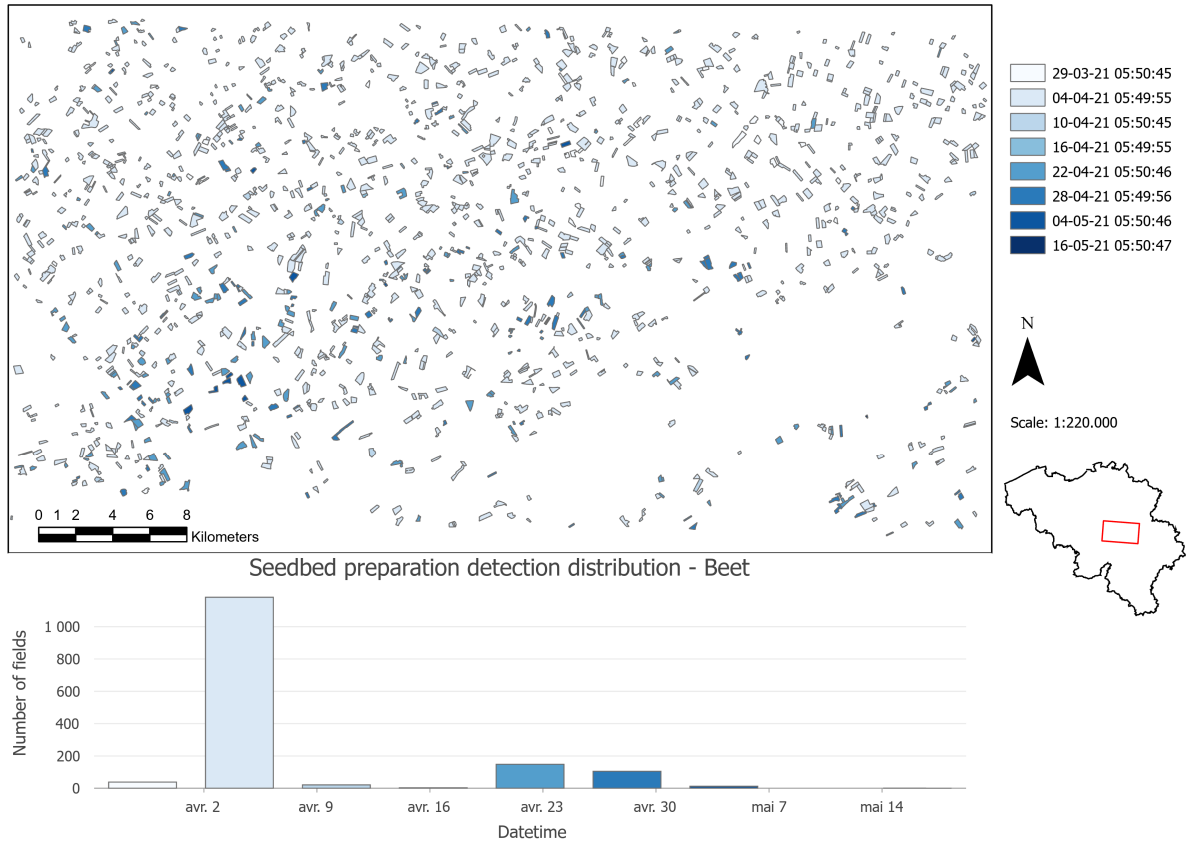


Figure 4.7: Spatial distribution map of the seedbed preparation date detected on the Sentinel-1 footprint for beet fields. The color intensity indicates the detected date with a stronger intensity being related to later dates. The bar plot shows the temporal distribution in number of fields, keeping the same color scheme.

Figure 4.8 showcases the same information for maize fields on the S1 footprint. The map shows no particular trend in distribution while the bar plot highlights some heterogeneity in the seedbed preparation date, being slightly concentrated around late April

and early May.

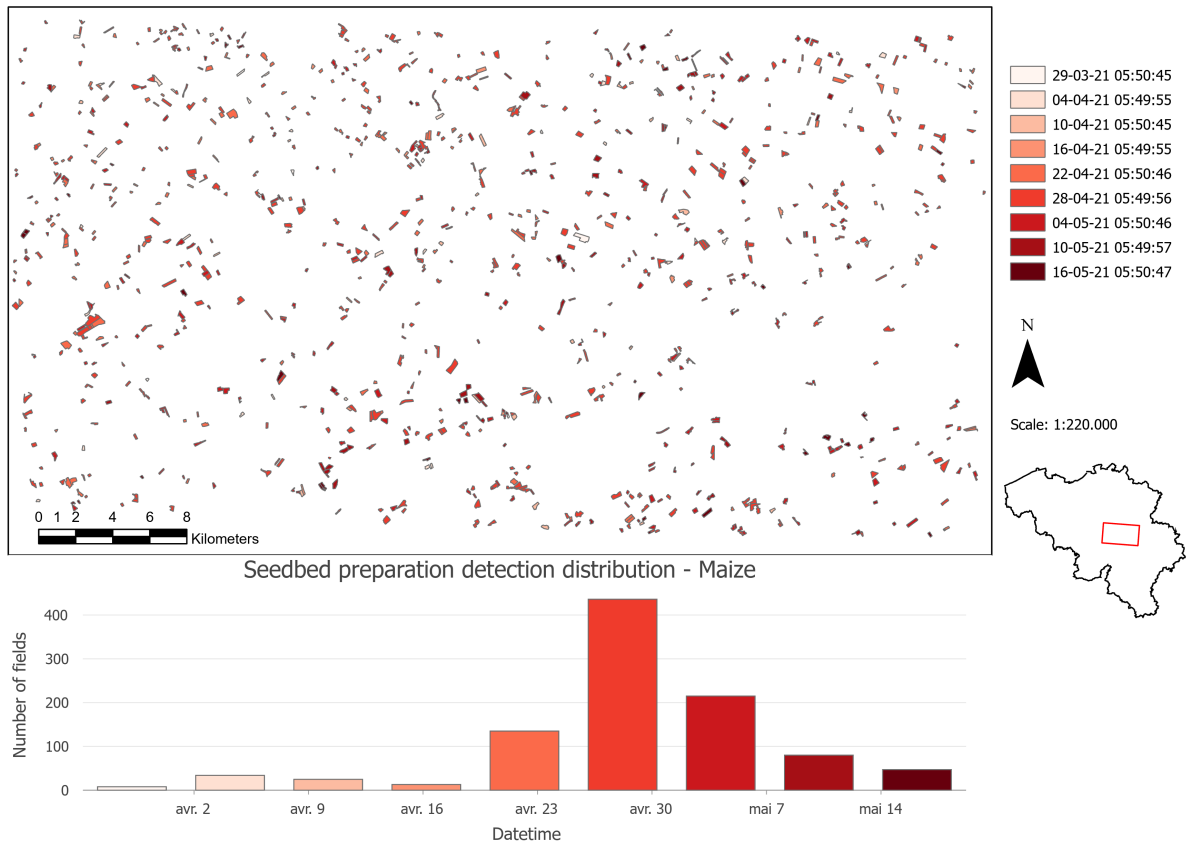


Figure 4.8: Spatial distribution map of the seedbed preparation date detected on the Sentinel-1 footprint for maize fields. The color intensity indicates the detected date with a stronger intensity being related to later dates. The bar plot shows the temporal distribution in number of fields, keeping the same color scheme.

Finally, Figure 4.9 exposes spatial and temporal heterogeneity in seedbed preparation detection for potato fields on the S1 footprint, with larger concentration on late April and early May.

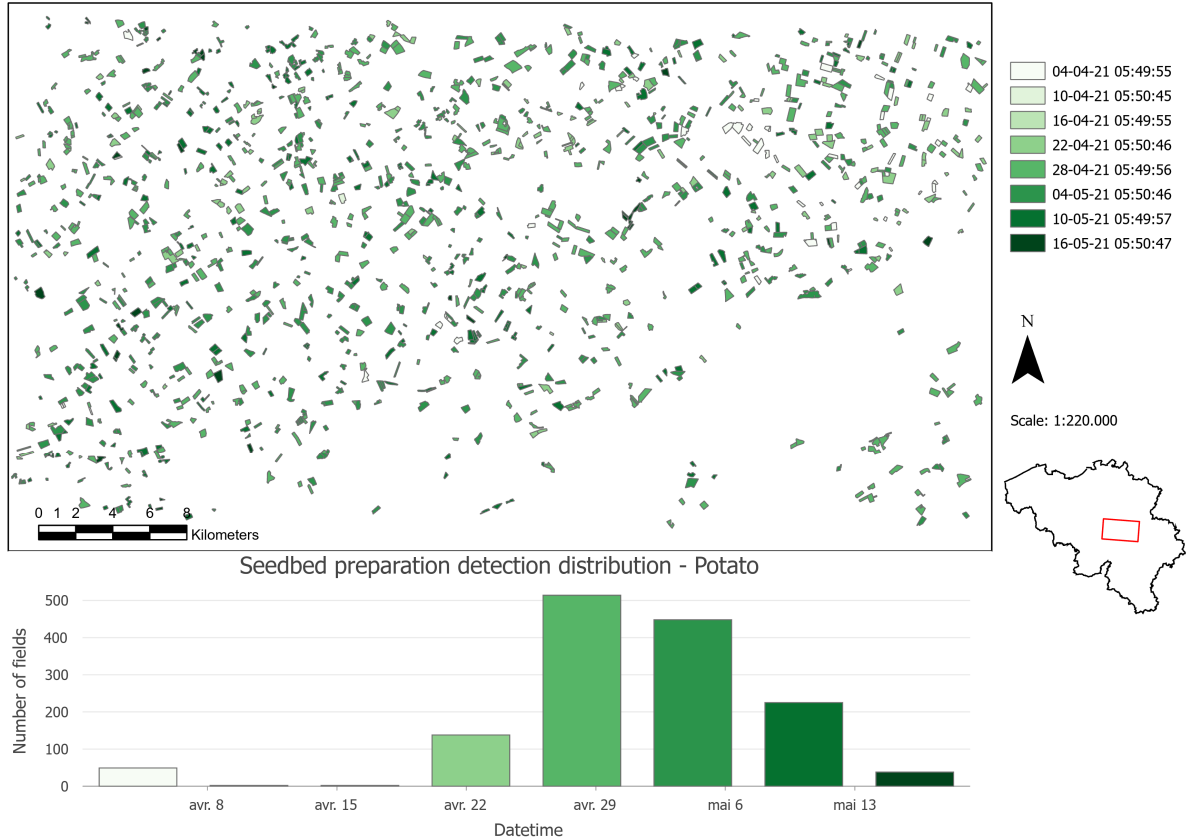


Figure 4.9: Spatial distribution map of the seedbed preparation date detected on the Sentinel-1 footprint for potato fields. The color intensity indicates the detected date with a stronger intensity being related to later dates. The bar plot shows the temporal distribution in number of fields, keeping the same color scheme.

4.2.3 Validation

The only available validation data found is extracted from the CRA-W agrometeorological report of the 2021 spring season (BCGMS, 2021). The report shows a histogram (Figure 4.10) of the temporal distribution in number of fields by observed sowing period for 162 beet fields in the Walloon region. In addition to that, a more detailed table retrieves some basic statistics for each agricultural region. Table 4.1 shows given statistics for the silty and Condroz regions (S1 footprint agricultural regions).

Table 4.1: Statistics of sowing dates (2021) for beet fields in the Silty region and the Condroz (BCGMS, 2021).

	Nb of fields	Mean	Min.	Max.
Silty region	78	04-05	03-24	04-23
Condroz	18	04-08	03-29	04-24

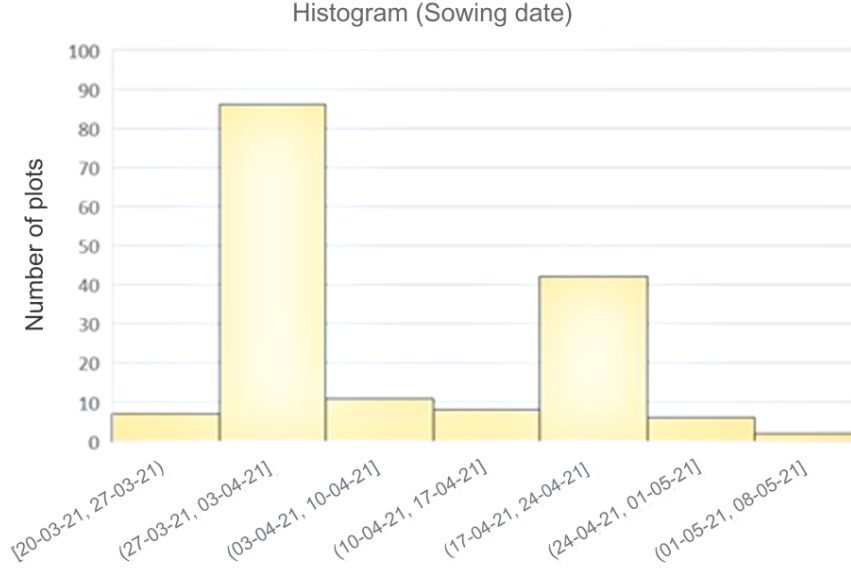


Figure 4.10: Histogram of beet sowing dates observed in the IRBAB network for 162 fields (BCGMS, 2021).

This histogram can be directly compared to the previously presented graph on the lower part of Figure 4.7 which also presents a peak during the first days of April and a secondary, much smaller peak around the 20th of April.

When applying the detection methodology to all fields visited during the campaign, an accuracy of 93.75% is achieved ($n = 32$) with the 20 m inner buffer and the minimum threshold of 50 pixels applied. The two fields for which the detected seedbed preparation date falls outside the observation period of the field campaign consistently show a minimum VV coherence value one acquisition later than expected. The wrongly detected fields are one beet field and one potato field. With the low amount of fields available and knowing the detection methodology has been developed using part of this data, it is hard to consider these accuracy results as reliable validation.

4.3 Early-season crop classification

The early-season crop classification constitutes the main focus of this study, emphasizing the usefulness of such data in assisting farmers with their crop management practices, as well as in facilitating fast decision-making and monitoring for both farmers and public institutions. This section presents Random Forest classification results using the F1 score metric with various scenarios of input data for the algorithm. An F1 score of 0.8 will serve as the minimum reference for a reliable classification. Moreover, feature importance is also plotted for the classification in the format of a MDI importance metric, highlighting which features are the most effective at dividing the input data into distinct classes.

In order to insure data representativeness and evaluate the impact of the criteria on classification, different minimum number of pixels per field thresholds are applied to both sensors acquisitions during the classification process. Table 4.2 gives the corresponding minimum surface for each pixel threshold value for both TerraSAR-X (3.5 m ground sampling distance) and Sentinel-1 (10 m ground sampling distance).

Table 4.2: Minimum surface covered in ha corresponding to every minimum pixel value threshold per field for TerraSAR-X and Sentinel-1.

	min. 50px	min. 100px	min. 150px
TerraSAR-X	min. 0.06 ha	min. 0.12 ha	min. 0.18 ha
Sentinel-1	min. 0.50 ha	min. 1.00 ha	min. 1.50 ha

As class distribution is not properly balanced (see Tables 4.3 & 4.4), SMOTE is applied to the training dataset in order to have the same amount of training samples per class for each Random Forest classification shown in the present section.

4.3.1 TerraSAR-X

Table 4.3 displays the class distribution for beet, maize, and potato fields for each minimum field size scenario (50, 100 and 150 minimum pixels per field) applied to TerraSAR-X images. As the threshold value increases, fewer fields are retained, and the remaining fields are the ones occupying larger areas.

Table 4.3: Class distribution for each crop type (beet, maize, potato) on the area of interest depending on the minimum image pixel threshold set by field for TerraSAR-X.

	no threshold	min. 50px	min. 100px	min. 150px
Beet	710	617	606	598
Maize	612	470	432	396
Potato	643	540	532	525

Figures 4.11, 4.12, and 4.13 display the evolution of the F1 score for Random Forest classification for each class and for respective threshold values of 50, 100, and 150 minimum pixels per field. The evolution is made with the cumulative addition of feature data depending on the acquisition date. An average F1 score is also plotted, showing the evolution of the mean F1 score for the classification.

While the general trend for each plot and crop type is towards an increase in the F1 score, notable differences can be observed between crops. For instance, beet fields tend to achieve a reliable score earlier than the other crop types, while maize fields often encounter

more challenges in being accurately classified earlier than June. Moreover, notable differences appear depending on the minimum pixels per field value, with the minimum of 100 pixels achieving best overall results for early classification of each crop type.

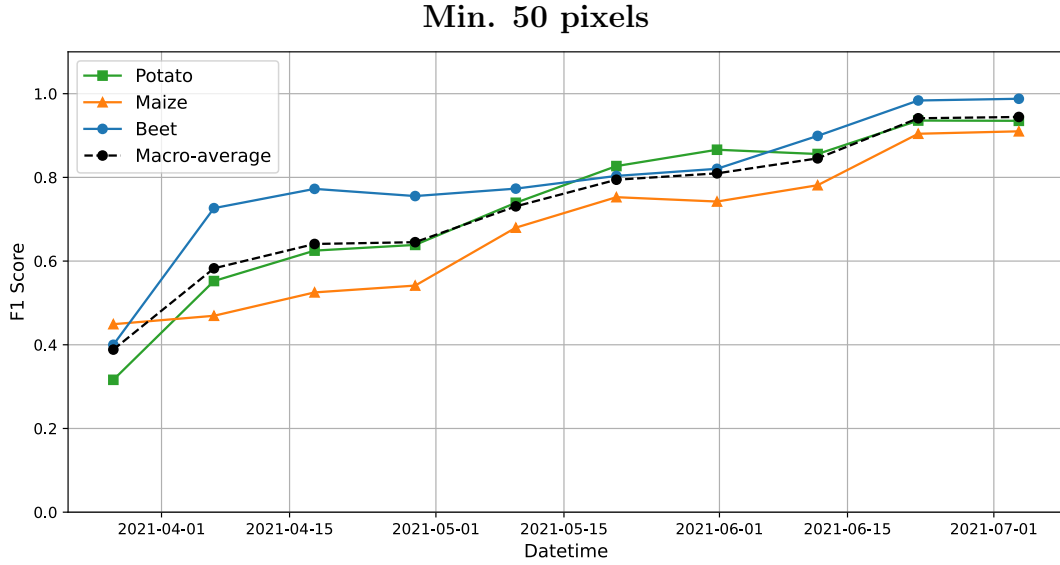


Figure 4.11: Evolution of F1 score for Random Forest classification of beet, maize, and potato crops on the area of interest with cumulative addition of TerraSAR-X acquisition data, while maintaining a minimum threshold of 50 pixels of image per field.

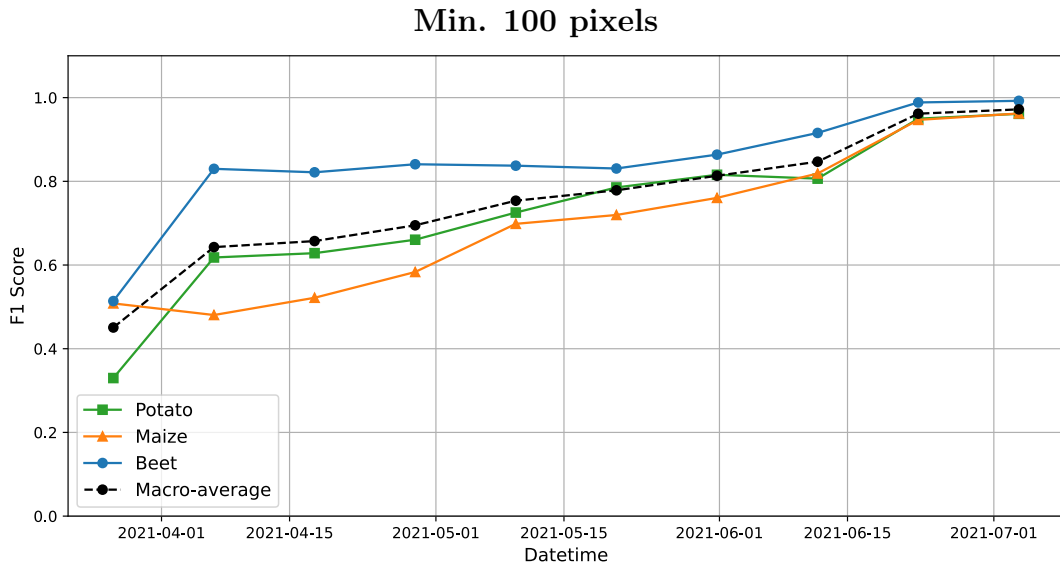


Figure 4.12: Evolution of F1 score for Random Forest classification of beet, maize, and potato crops on the area of interest with cumulative addition of TerraSAR-X acquisition data, while maintaining a minimum threshold of 100 pixels of image per field.

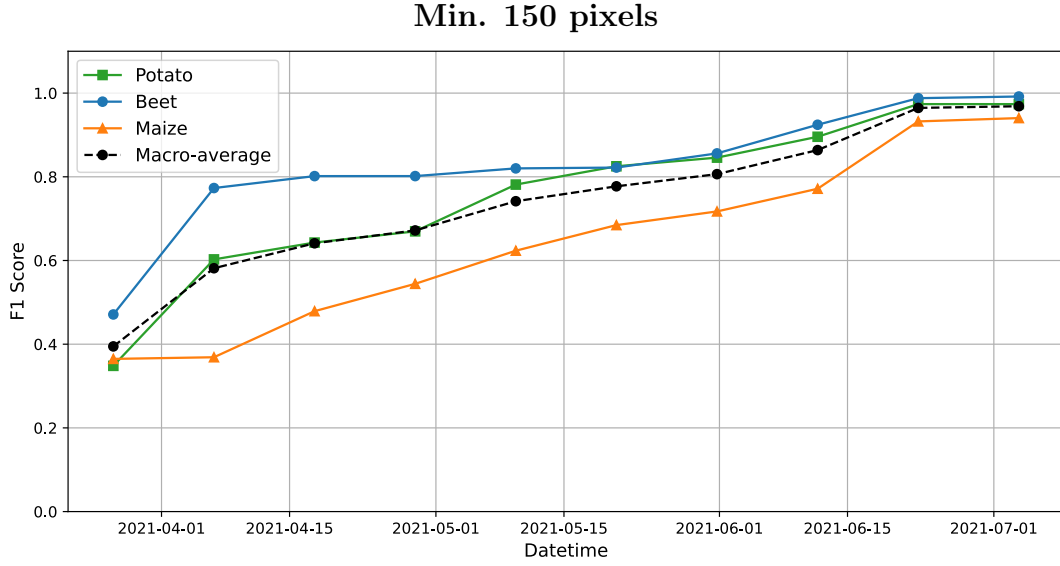


Figure 4.13: Evolution of F1 score for Random Forest classification of beet, maize, and potato crops on the area of interest with cumulative addition of TerraSAR-X acquisition data, while maintaining a minimum threshold of 150 pixels of image per field.

Figure 4.14 displays the top ten most important features contributing to the classification process in the scenario of a minimum pixel threshold of 50. VH and VV mean backscatter emerge as the most influential features for this classification. Additionally, the plot provides insights into the temporal aspect of each feature, clearly highlighting that the most important features for classification are derived from the data of the two most recent acquisitions.

Feature importance plots for the two other threshold values have similar hierarchy with some slight changes in MDI importance index values. These plots are available in Appendix B, Figures 8.20 & 8.21.

Finally, Figure 4.15 depicts F1 score evolution plots for each crop type with a minimum pixel threshold value of 50 pixels of TerraSAR-X image per field. Each plot results from a different classification, where the input data progressively exclude the earliest acquisitions. For example, the upper left plot displays the classification as detailed in Figure 4.11, while the upper right plot uses the same methodology but removes the first acquisition. These results provide insights into the potential importance (or lack thereof) of the initial acquisitions in the classification process. In this case, removing the first image has little impact on the F1 score results and even slightly improves the scores for maize and beet fields in late April. Conversely, classifications starting from the third and fourth acquisitions (mid- and late April) significantly affect the F1 score for every crop type in the early season. Every score drops by at least 0.1 compared to the classification using the complete acquisition data (upper left).

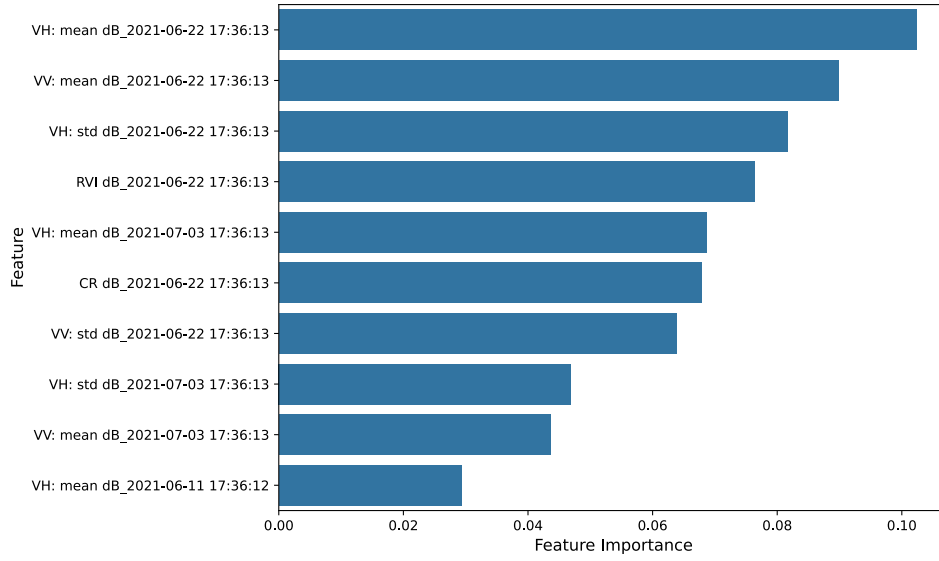


Figure 4.14: Top ten of most important features for the Random Forest classification of beet, maize, and potato crops using TerraSAR-X data on the area of interest, while maintaining a minimum threshold of 50 pixels of image per field.

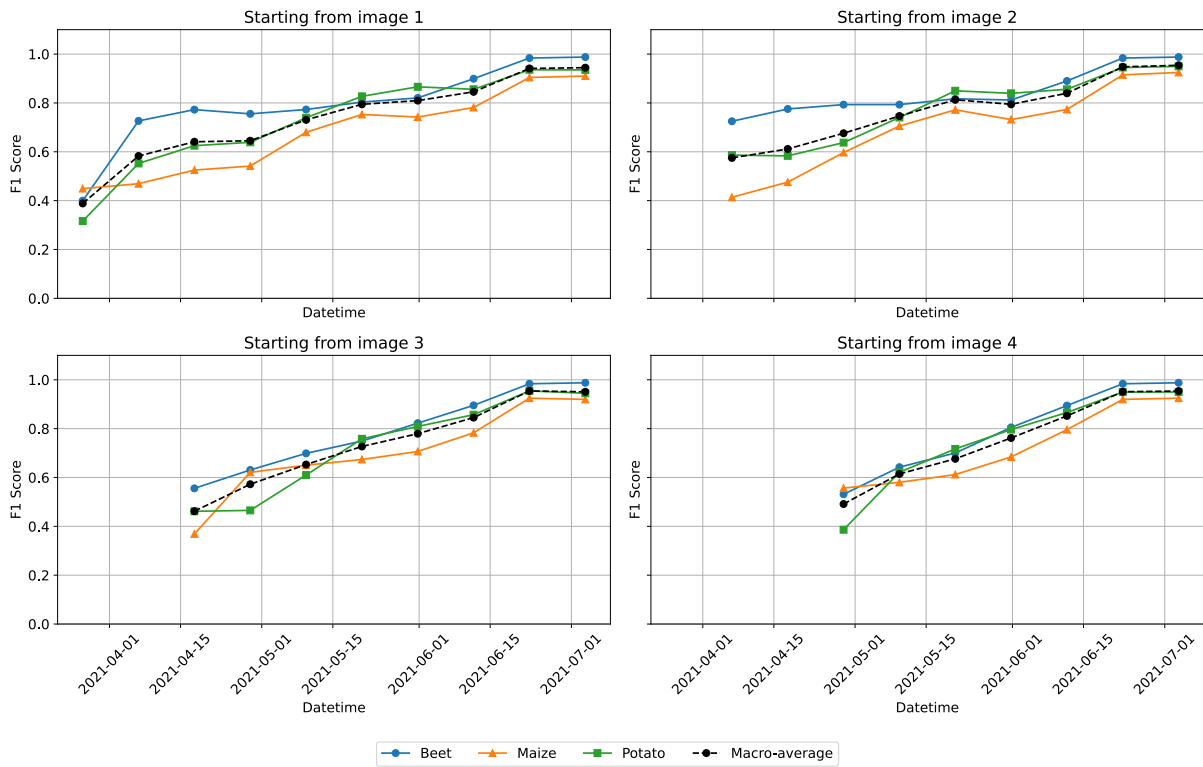


Figure 4.15: F1 score evolutions for the Random Forest classifications of beet, maize, and potato crops on the area of interest with cumulative addition of TerraSAR-X acquisition data. Each plot shows classification performance where the input data progressively exclude the earliest acquisitions. A minimum threshold of 50 pixels of image per field is applied.

4.3.2 Sentinel-1

Area of interest (AOI)

Similarly to the previous detailed process for TerraSAR-X, Table 4.4 presents the class distribution for beet, maize, and potato fields for each minimum pixel threshold scenario (50, 100, and 150 minimum pixels per field) applied to Sentinel-1 images. As the threshold value increases, fewer fields are retained, and the remaining fields are the ones covering larger areas.

Table 4.4: Class distribution for each crop type (beet, maize, potato) on the area of interest depending on the minimum image pixel threshold set by field for Sentinel-1.

	no threshold	min. 50px	min. 100px	min. 150px
Beet	710	567	476	405
Maize	612	325	204	152
Potato	643	531	457	398

Figures 4.16 to 4.18 provide a visual representation of the F1 score evolution for each studied crop type, along with the associated average F1 score evolution for 50, 100, and 150 minimum pixels of Sentinel-1 image per field.

Overall, potato and beet fields experience a significant increase in F1 score in early April, while maize fields tend to plateau and even exhibit notable declines during the month, indicating challenges in achieving accurate classification quickly in the season.

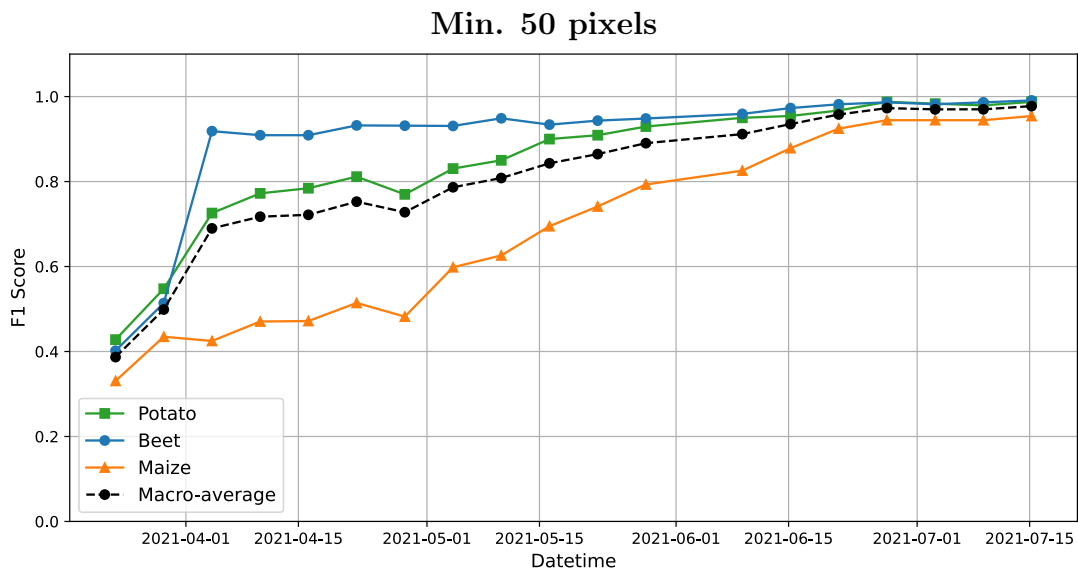


Figure 4.16: Evolution of F1 score for Random Forest classification of beet, maize, and potato crops on the area of interest with cumulative addition of Sentinel-1 acquisition data, while maintaining a minimum threshold of 50 pixels of image per field.

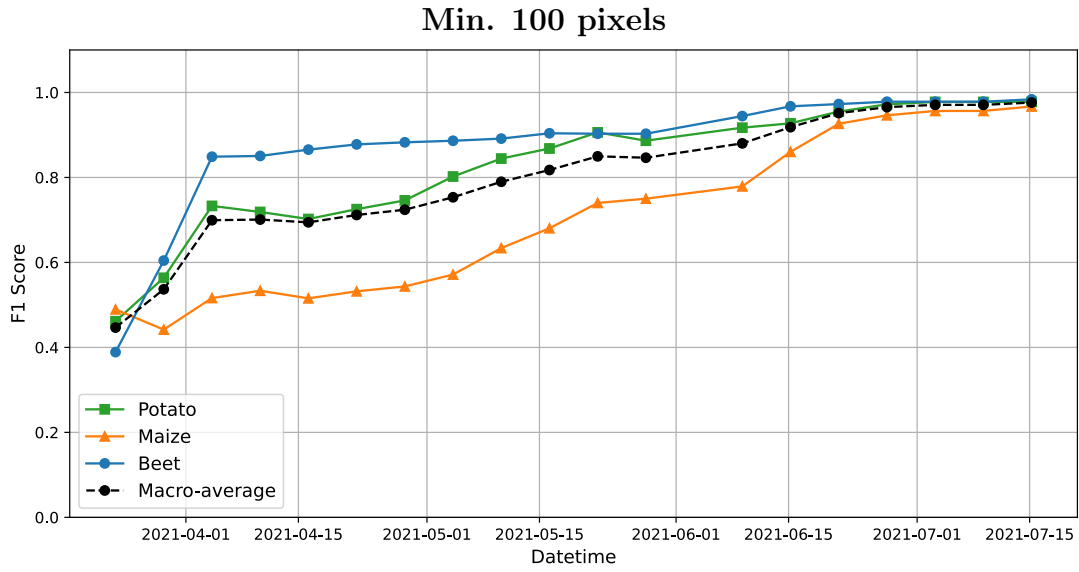


Figure 4.17: Evolution of F1 score for Random Forest classification of beet, maize, and potato crops on the area of interest with cumulative addition of Sentinel-1 acquisition data, while maintaining a minimum threshold of 100 pixels of image per field.

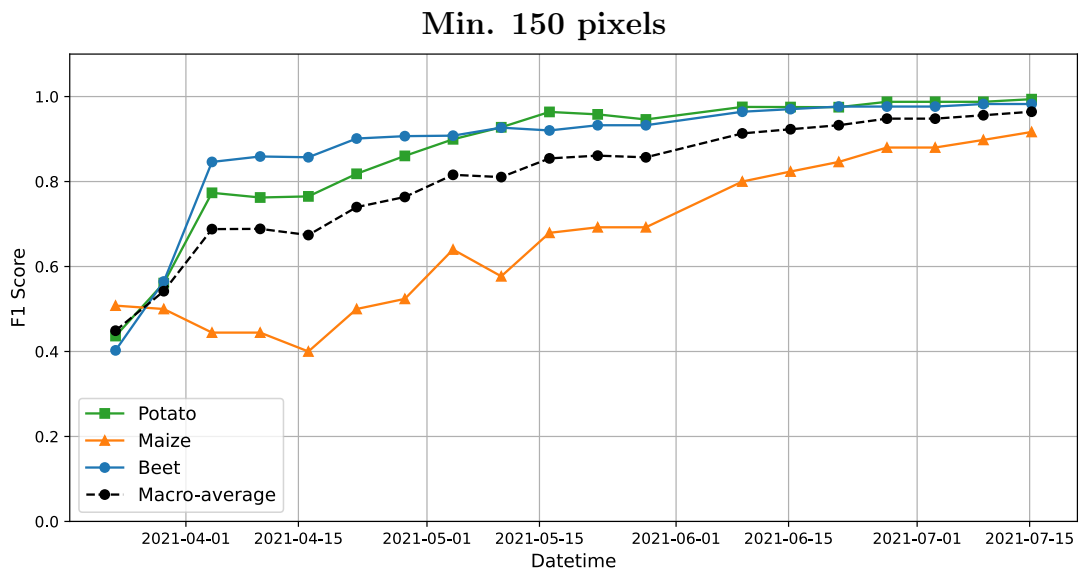


Figure 4.18: Evolution of F1 score for Random Forest classification of beet, maize, and potato crops on the area of interest with cumulative addition of Sentinel-1 acquisition data, while maintaining a minimum threshold of 150 pixels of image per field.

The most important features for the Random Forest classification using Sentinel-1 C-band data are VV and VH backscatter related features with mean VV coherence also having significant importance in the accurate classification of each crop type. Again, important features are also features derived from late season data. Figure 4.19 shows the top ten most important features in the case of a 50 minimum pixels threshold value. Plots for threshold values of 100 and 150 pixels are available in Appendix B, Figures 8.22 & 8.23.

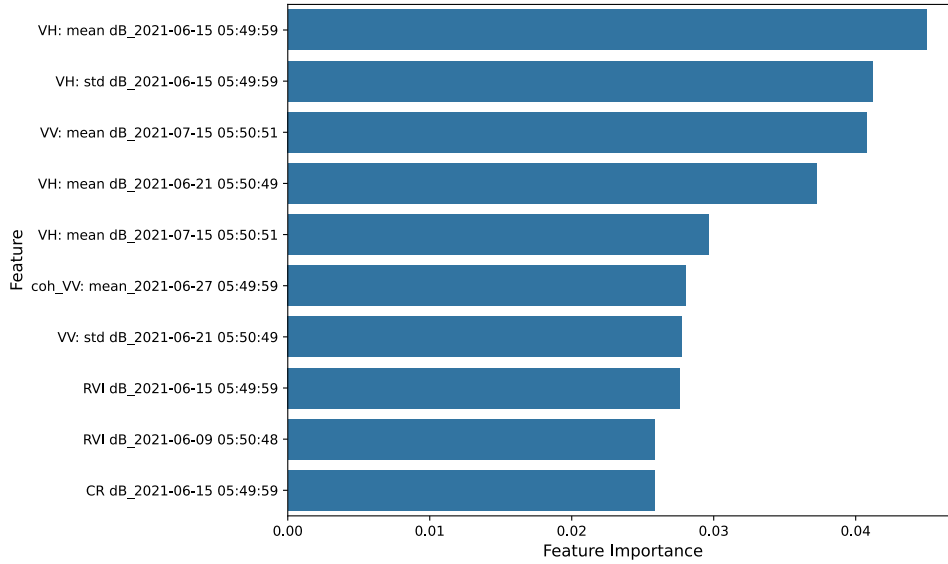


Figure 4.19: Top ten of most important features for the Random Forest classification of beet, maize, and potato crops using Sentinel-1 data on the area of interest, while maintaining a minimum threshold of 50 pixels of image per field.

Figure 4.20 shows F1 score evolutions for different classifications, using different sets of Sentinel-1 acquisitions input data and progressively excluding the first images of the dataset. Even though slight changes in F1 score appear when removing the first four images, significant impact is only noticed when starting from the sixth image (22nd of April). Indeed, F1 scores for beet and potato fields are strongly impacted, losing around a tenth compared to F1 scores at the same time using the complete dataset of acquisitions.

Complete Sentinel-1 footprint

Given that the area defined by the S1 footprint encompasses a wide distribution of fields in the silty region (see Figure 3.1 in Section 3.1), an extended classification has been applied to all fields within the footprint to evaluate the significance of data quantity for the training process of the classification algorithm. The class distribution within the complete Sentinel-1 footprint comprises 1510 beet fields, 993 maize fields, and 1416 potato fields, following the 50 minimum pixels per field threshold. After applying an 80-20 split for training and testing samples and using SMOTE on the training dataset, each class in the new training sample comprises 1208 samples.

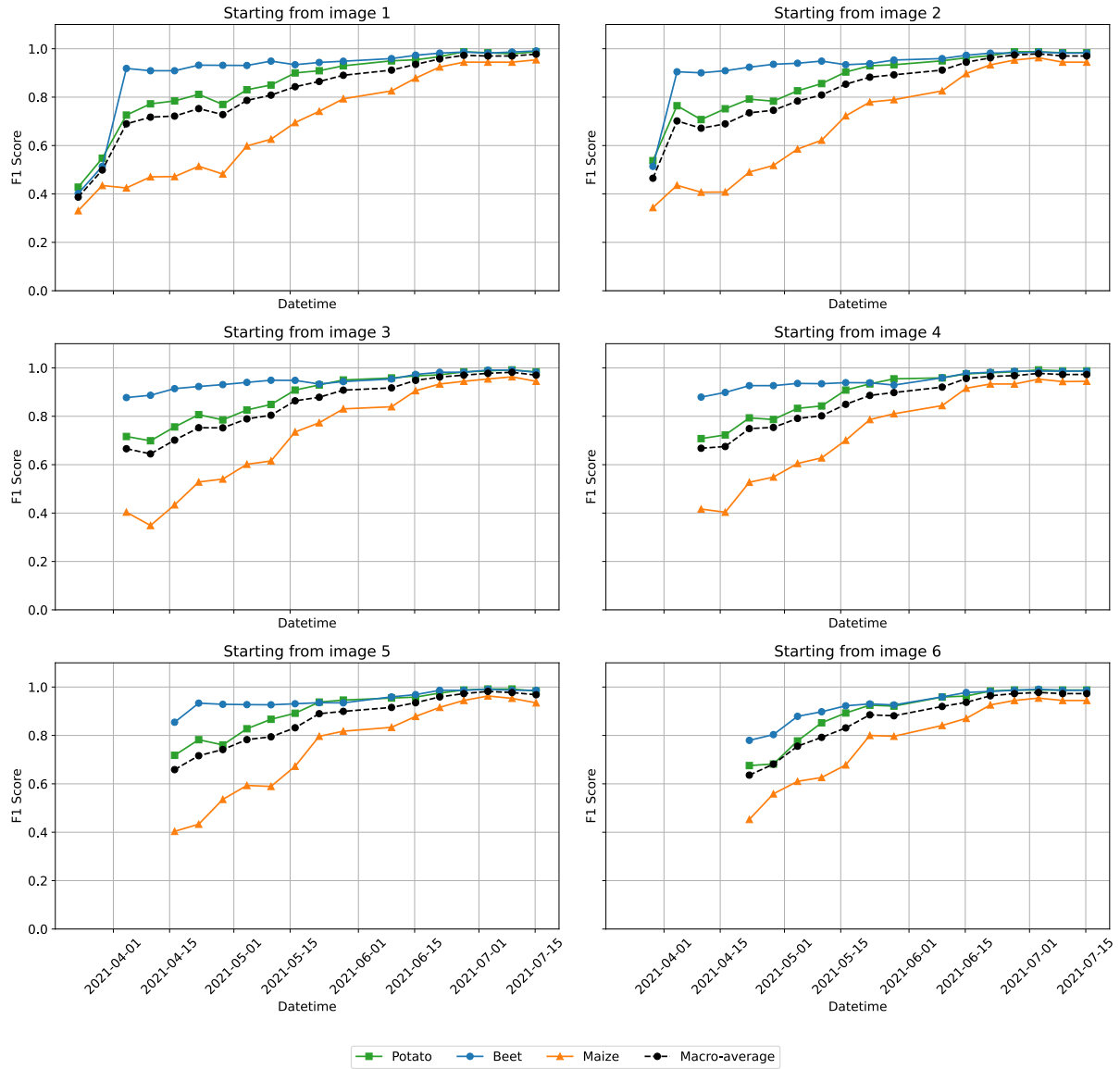


Figure 4.20: F1 score evolutions for the Random Forest classifications of beet, maize, and potato crops on the area of interest with cumulative addition of Sentinel-1 acquisition data. Each plot shows classification performance where the input data progressively exclude the earliest acquisitions. A minimum threshold of 50 pixels of image per field is applied.

Figure 4.21 illustrates the evolution of the F1 score derived from the Random Forest classification with the cumulative addition of Sentinel-1 data covering the entire footprint. Beet fields achieve sufficiently accurate classification (F1 score above 0.8) less than a week after the mean sowing date for beet fields on this same area, the F1 score then stabilizes until late May. Maize and potato fields exhibit a less pronounced increase in F1 score in the early season, with a slower rate of increase over time. An almost constant divergence is observed for each crop type until late May or early June, where F1 scores appear to converge.

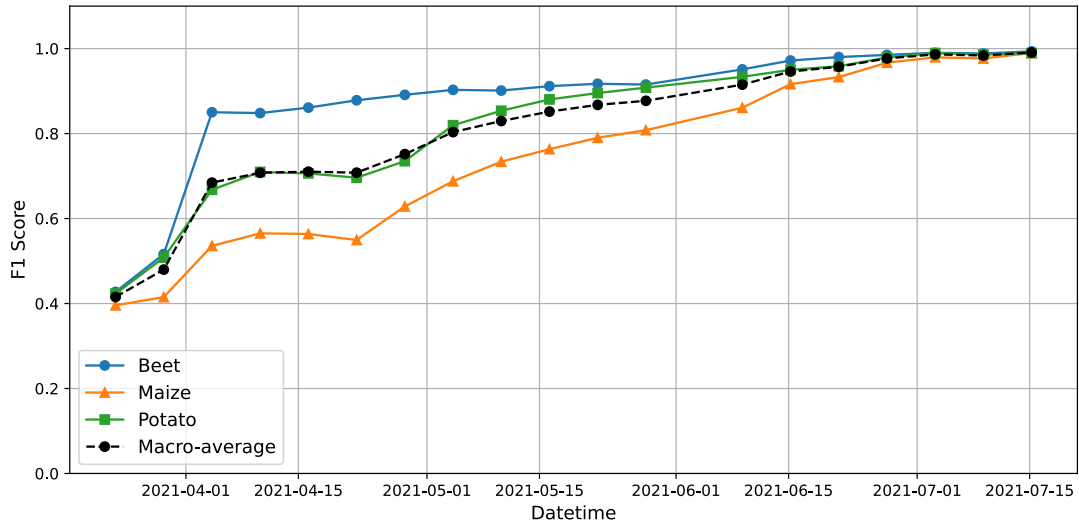


Figure 4.21: Evolution of F1 score for Random Forest classification of beet, maize, and potato crops on the complete Sentinel-1 footprint with cumulative addition of Sentinel-1 acquisition data, while maintaining a minimum threshold of 50 pixels of image per field.

4.3.3 Sentinel-1 reduced frequency for comparison purposes

In order to compare Sentinel-1 and TerraSAR-X classifications properly, it is necessary to compare them at the same level. This is why a virtual reduction of revisit period is applied on Figure 4.22, resulting in a total of 10 Sentinel-1 acquisitions instead of 20 with a fake revisit period of 12 days (as if only either S-1A or S-1B were available), almost matching the TerraSAR-X revisit period.

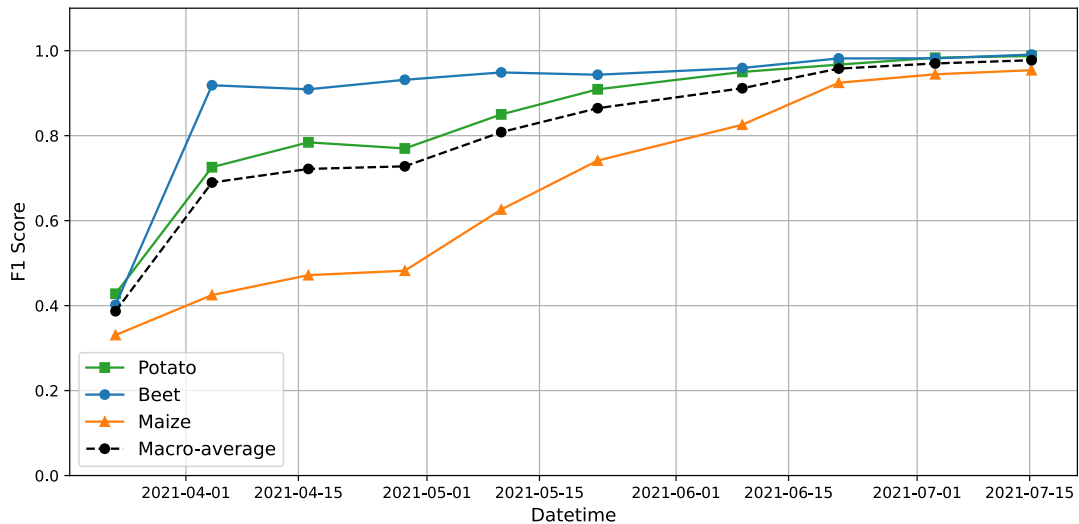


Figure 4.22: Evolution of F1 score for Random Forest classification of beet, maize, and potato crops on the area of interest with cumulative addition of a reduced amount of 10 Sentinel-1 acquisitions data, while maintaining a minimum threshold of 50 pixels of image per field.

4.3.4 TerraSAR-X and Sentinel-1 data combination

Finally, the combination of both TerraSAR-X and Sentinel-1 data is displayed on Figure 4.23 to evaluate how it performs. Again, only the closest Sentinel-1 acquisition to each one of the TerraSAR-X acquisitions is kept to combine both data for the same period of time and create a total of 16 features (10 for S1 and 6 for TSX) per field for each acquisition date.

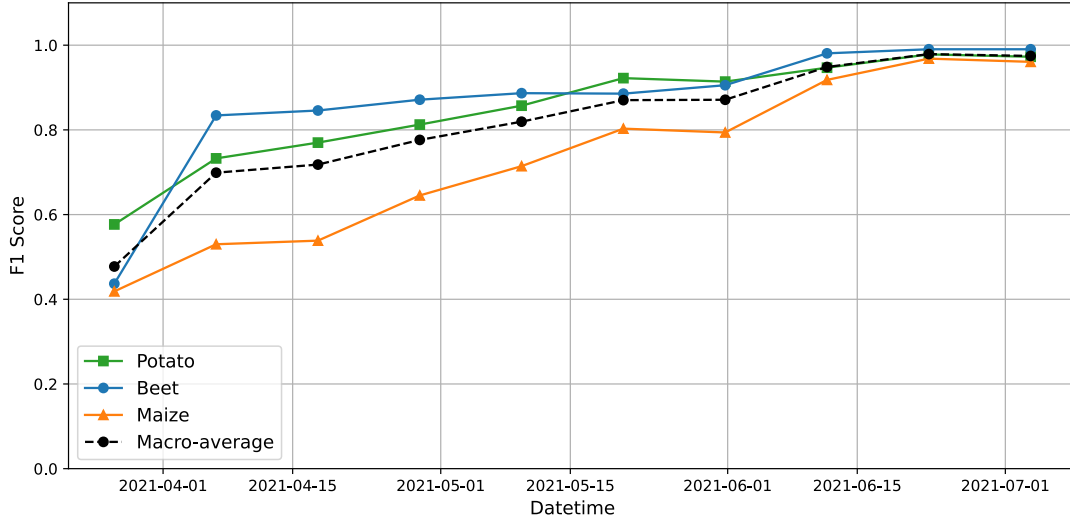


Figure 4.23: Evolution of F1 score for Random Forest classification of beet, maize, and potato crops on the area of interest with cumulative addition of a combination of TerraSAR-X and Sentinel-1 acquisitions data, while maintaining a minimum threshold of 50 pixels of image per field.

Figure 4.24 reveals that the top nine out of the ten most important features in the classification, resulting from the combination of TerraSAR-X and Sentinel-1 data, are related to TerraSAR-X data. Only the tenth most important feature is associated with Sentinel-1 data. Additionally, all of the top ten features are also referring to features derived from the last acquisitions data.

Figure 4.25 shows different F1 score evolutions using the combination of Sentinel-1 and TerraSAR-X acquisitions derived data. Each plot is the result of a different Random Forest classification process, with progressive removal of the first acquisitions data. Some notable impacts are observed on F1 scores of beet and potato fields with a classification starting from the fourth acquisition, in late April, when compared to the initial classification (upper left plot).

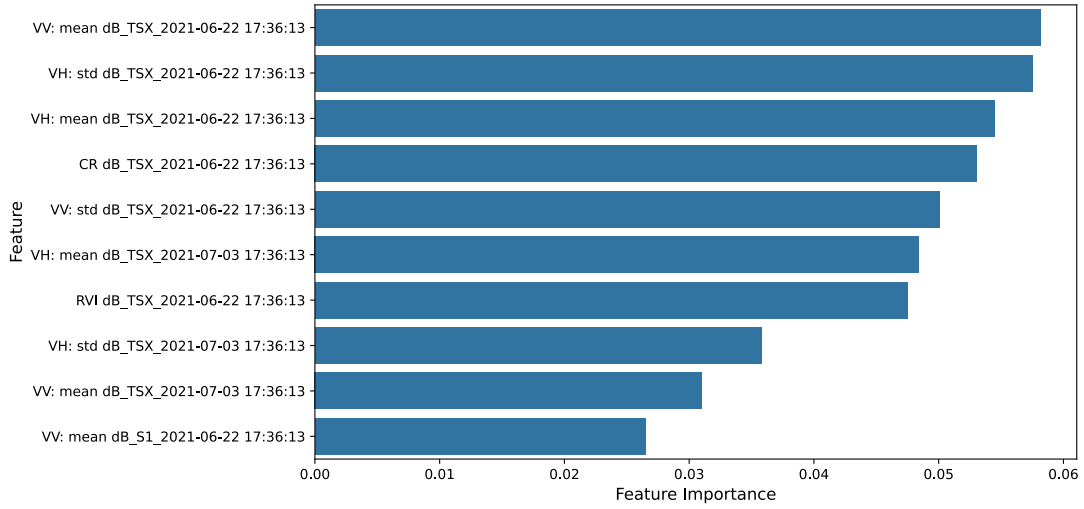


Figure 4.24: Top ten of most important features for the Random Forest classification of beet, maize, and potato crops using Sentinel-1 and TerraSAR-X data on the area of interest, while maintaining a minimum threshold of 50 pixels of image per field.

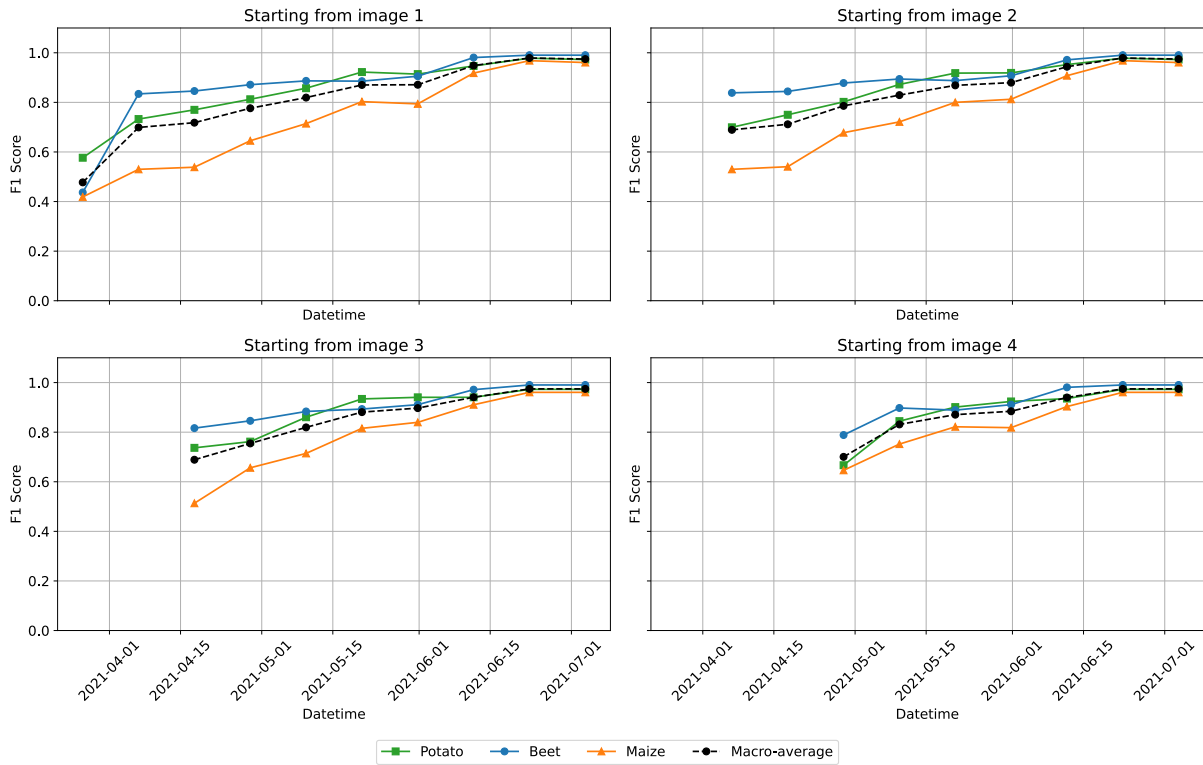


Figure 4.25: F1 score evolutions for the Random Forest classifications of beet, maize, and potato crops on the area of interest with cumulative addition of TerraSAR-X and Sentinel-1 acquisitions data. Each plot shows classification performance where the input data progressively exclude the earliest acquisitions. A minimum threshold of 50 pixels of image per field is applied.

Finally, in order to understand how the final Random Forest classification operates, it is insightful to render confusion matrices, which visualize the predicted labels for each class of crop type. Figure 4.26 displays the first four confusion matrices of the combined

classification with Sentinel-1 and TerraSAR-X data, employing the usual minimum 50 pixels threshold value.

The first matrix (Figure 8.24a) indicates a tendency to misclassify beet and maize fields as potato fields, or to classify beet fields as maize fields. This trend quickly diminishes in the second matrix (Figure 8.24b), where a substantial number of beet fields are correctly classified, although maize fields still suffer from misclassification as potato fields. The overall trend fades throughout the classification process to achieve an almost perfect classification in the end. The remaining six matrices are available in the Appendix B, Figure 8.24.

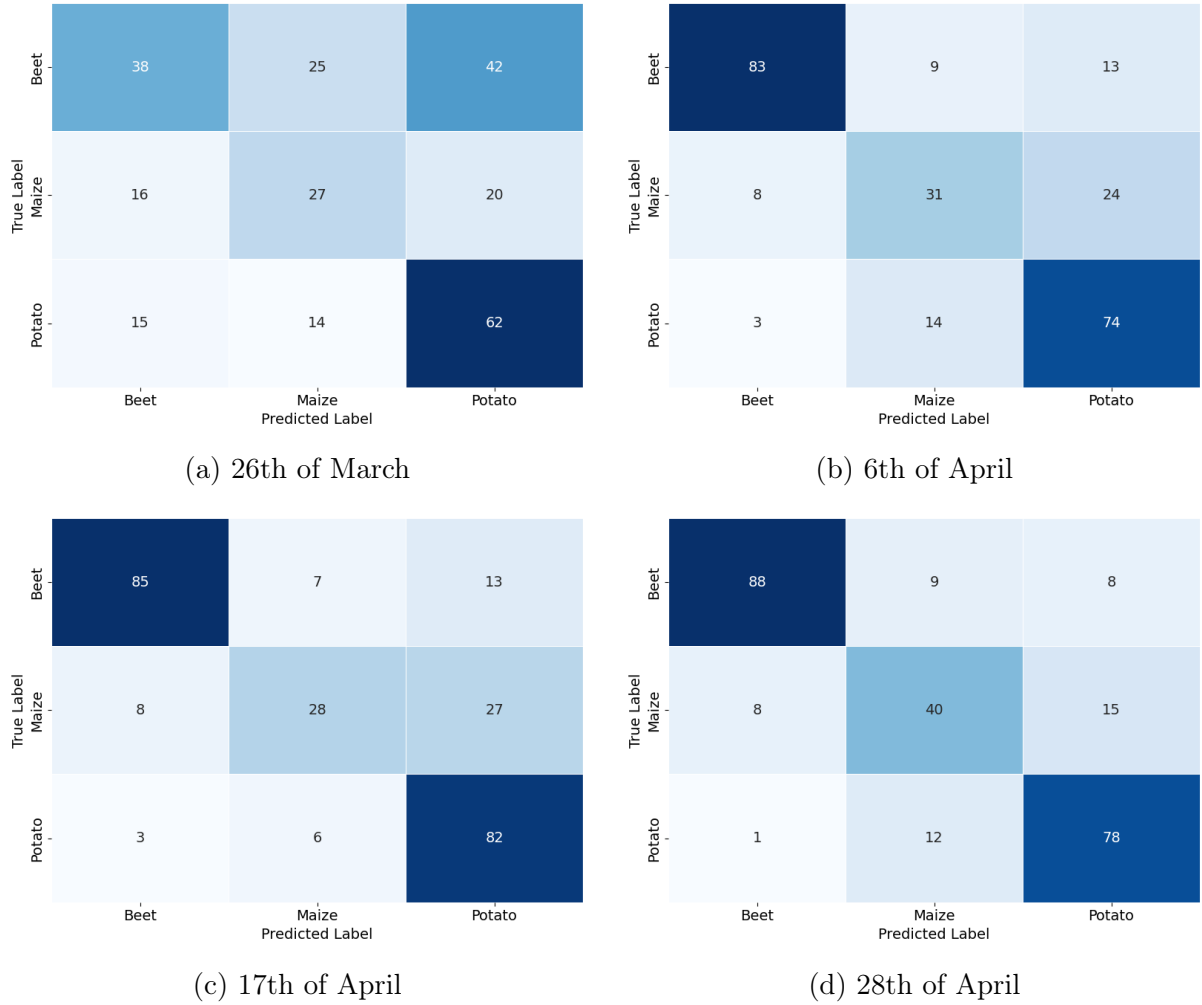


Figure 4.26: Four first confusion matrices derived from the Random Forest classification of beet, maize, and potato crops on the area of interest with cumulative addition of a combination of TerraSAR-X and Sentinel-1 acquisitions data, while maintaining a minimum threshold of 50 pixels of image per field.

4.3.5 Classification accuracy: summary

To summarize the computed results, Table 4.5 outlines the earliest acquisition date at which the F1 score of each classification scenario reaches 0.8, a threshold arbitrarily set to describe reliable classification.

Table 4.5: Earliest acquisition date at which the F1 score of each Random Forest classification process for beet, maize and potato fields on the area of interest reached 0.8. Different scenarios are considered for TerraSAR-X (TSX) and Sentinel-1 (S1) with varying thresholds of minimum pixels covering each field. Color intensity related to the acquisition date is darker when the score is reached earlier in the season and lighter when the score is reached later in the season.

	Beet	Maize	Potato
TSX min. 50px	05-20	06-22	05-20
TSX min. 100px	04-06	06-11	05-31
TSX min. 150px	04-17	06-22	05-20
S1 min. 50px	04-04	06-09	04-22
S1 min. 100px	04-04	06-15	05-04
S1 min. 150px	04-04	06-15	04-22
S1 10 acquisitions	04-04	06-09	05-10
TSX and S1	04-06	05-20	04-28

DISCUSSION

This chapter discusses the previously presented results, providing a critical analysis for the two main objectives of the study. Additionally, a section is dedicated to the main limitations and future perspectives of the studied subject.

5.1 Seedbed preparation detection with InSAR coherence time series

Recent agrometeorological reports provided by the CRA-W and the RMI have proven the importance and the challenge of providing accurate sowing dates in the early season. As of mid-May 2024, the CIPF (Walloon advisory service for maize) estimated that only 30% of the maize fields were sown, emphasizing the difficulty of providing a precise quantification of sown surface under varying wet weather conditions (BCGMS, 2024).

The study results unfortunately demonstrate that the long revisit period and the lack of coherence computation capabilities of TerraSAR-X are the two main reasons why seedbed preparation detection has been challenging with the available X-band data. Time series of VV and VH backscatter values fail to reveal a precise and clear enough trend in the signal to identify a seedbed preparation phase that aligns with the sowing period observed in the field campaign data.

On the other hand, C-band data obtained from Sentinel-1 with a close revisit period enables a more precise visualization of temporal dynamics compared to TerraSAR-X. It clearly highlights drops in VV and VH backscatter, as well as VV and VH coherence, matching the field campaign sowing period. This is particularly evident in the VV coherence time series, where coherence values rapidly transition from high to low and then back to high again, indicating a significant change in soil properties (here, roughness and moisture content) from a period of time to another. Based on these preliminary results,

a strong hypothesis is formed that the seedbed preparation date can be detected by identifying the acquisition date for which the VV coherence value is the lowest between the 15th of March and the 31st of May (considered as the seedbed preparation period).

The full extension to every beet, maize, and potato field covered by the selected Sentinel-1 footprint reveals the spatial and temporal distribution of the detected date for each crop type. However, the lack of reliable validation data makes it challenging to definitively evaluate the accuracy of the chosen detection method, although the BCGMS report plots and statistics appear to align with the results outputs for beet fields. Applying the process to all visited fields during the campaign returns an overall accuracy over 93% but the reliability of this result is still compromised mainly due to the small sample size.

It is also essential to highlight that the detected "date" actually corresponds to a Sentinel-1 acquisition date at which the VV coherence is at its lowest within a fixed time period. In reality, the seedbed preparation occurs between the previous acquisition date and the detected one, resulting in a detected seedbed preparation period rather than a precise date. This underscores the importance of a close revisit period for the detection of seedbed preparation, in order to achieve a very narrow period detection.

5.2 Early-season crop classification

The early-season crop classifications show promising results, comparing the use of different radar sensors and their respective characteristics alongside with different pre-processing parameters.

5.2.1 Impact of the minimum pixel threshold on classification results

Firstly, TerraSAR-X and Sentinel-1 present very different class distributions for each crop type and minimum pixel value due to their respective GSD. Indeed, since TerraSAR-X allows a finer GSD compared to Sentinel-1, more fields can be kept under the minimum pixel per field constraint for each crop type. This leads to a greater discrimination in fields for Sentinel-1 C-band data, therefore impacting the total size of the training sample of the model and consequently, output results. When comparing TerraSAR-X and Sentinel-1 classification results, the 150 and 50 minimum pixel scenarios will be considered respectively for TerraSAR-X and Sentinel-1 as they are the closest in terms of class distribution. Let's also note that since maize fields are generally smaller fields compared to potato and beet fields, the minimum pixel per field constraint impacts this crop type

even more in class distribution when increasing the threshold value.

The impact of the minimum pixel threshold for every field can be significant depending on the sensor. There is a clear trade-off between applying a higher threshold for better signal-to-noise ratio for each crop type and reducing the threshold value to include a larger number of fields for the training of the model.

For example, the difference between the 50 and 100 threshold value for TerraSAR-X can be seen as an increase in representativeness rather than a downside effect to shrinking training data since all three crop types globally perform better through the evolution of their F1 score. The 150 pixel threshold also leads to some global better results but penalizes maize in the early season due to a lower number of fields represented in the data, which goes in the opposite direction of the fixed objective of early classification.

On the other hand, Sentinel-1 data shows a slightly different reaction to the pixel threshold value. By again comparing the 50 and 100 pixel thresholds, a clear downshift in potato and beet field classification performance is observed in the early season while maize fields presents slightly higher F1 scores during that same period. The 150 pixel threshold F1 score evolution clearly highlights the impact of unbalanced data for classification, showing strong variations in F1 score for maize (152 fields) while potato and beet fields (respectively 398 and 405 fields) classify better than the two first examples.

Even though SMOTE is applied to the training dataset to mitigate the effects of imbalance, it does not totally correct it. Indeed, oversampling can sometimes amplify noise that can be present in the initial dataset and the quality of synthetic samples might also vary, especially for underrepresented classes such as maize fields in this case.

Overall, the major impact on classifications will be the imbalance between classes in the input dataset. A clear trade-off exists between signal quality and the number of fields of each class, showing varying results depending on the amount of pixels covering each field and depending on the sensor’s ground sampling distance. Results based on the complete Sentinel-1 footprint have also shown that the input dataset size is of great importance, allowing a larger training dataset.

5.2.2 X-band and C-band: respective contributions and comparison

Previous results highlighting the seedbed preparation period for each crop type have shown that a large majority of beet fields have their fields prepared for seeding in the first days of April while maize and potato fields are mostly prepared in the late days of April

or begin May. These periods of time can strongly be related to sudden increases in F1 score when looking at classification results. For example, beet fields are quickly classified accurately in early April while F1 score increases more rapidly in the first days of May for potato and maize fields. Based on this, it seems important to put attention to these sowing dates when comparing classification results in time. It is still important to remember that even though almost all beet fields are sown during the same period of time, it isn't the case for potato and maize fields which present a much more heterogeneous distribution in sowing dates. The mean sowing period for each crop type will serve as reference.

While TerraSAR-X performs well for the classification of maize fields on the overall early season for a pixel threshold value of 50, its performance deteriorates as this value increases to 150 (TSX threshold value for comparison with S1). In that case, its classification performance aligns more closely with the Sentinel-1 F1 score evolution, reaching a score of 0.6 in early May and 0.8 in mid-June, about two weeks after the mean sowing date of maize fields. Concerning potato fields, Sentinel-1 reaches a plateau nearing an F1 score of 0.8 during the early stages of April, roughly two weeks before the mean sowing date while TerraSAR-X results exhibit a similar plateau approximately a month later in the season.

These particular differences between sensors could be explained by the physical properties of C-band and X-band. Indeed, C-band uses a lower frequency signal allowing better sensitivity to sub-surface soil layers (up to several tens of centimeters into the ground) while X-band data is sensitive to the monitoring of smaller-scale surface features. Potato and maize fields are managed very differently, with the formation of hills on potato fields while maize parcels can present much smoother surfaces, playing crucial role in classification of both crop types.

Feature importance results converge for both C-band and X-band, commonly stating that VV & VH backscatter play significant roles in accurate classification but more importantly that the most discriminant features are the ones extracted from the latest acquisitions. These final acquisition dates in late June and early July coincide with significant vegetation growth above the soil, which varies by crop type. This variation impacts the backscattering signal, potentially explaining the importance of these acquisitions in the classification process.

Finally, as the temporal resolution of Sentinel-1 allows finer details in time series and consequently F1 score evolution visualization, a virtual reduction of the revisit time is applied by simulating a revisit period of 12 days to match TerraSAR-X's 11 days revisit period as closely as possible. When comparing Sentinel-1 F1 score evolution with the actual revisit period of 6 days and the simulated 12 days revisit period, some important

plateaus are reached a few days earlier in the season due to more precise temporal information. For example, maize fields reach the 0.8 F1 score level slightly before June, while this plateau is reached a few days later when the temporal resolution is doubled. This emphasizes the importance of temporal resolution in a situation like agriculture assistance where having precise information a few days earlier can be of great importance.

5.2.3 Combination of X-band and C-band: multi-frequency as a solution?

TerraSAR-X and Sentinel-1 have respectively revealed some opposite advantages and disadvantages for the three crop types classification. But what if combining both of their data could be a possible solution to extract the best of X-band and C-band information ?

Results indicate that combining the fine surface information from X-band data with the sub-surface information from C-band data yields the best results per class for early-season classification. Beet fields surpass the 0.8 F1 score value in early April, only days after the mean sowing period, maize fields achieve it in mid to late May, two to three weeks after the mean sowing period and potato fields reach it during the last days of April, roughly matching the mean sowing period of that culture.

Feature importance analysis highlights TerraSAR-X data as the most discriminant sensor for the accurate classification. However, it's important to note that these feature importance results are related to the global classification and primarily focus on the last acquisitions for the most important features of the classification.

All of this highlights the importance of multi-frequency data for classification purposes, combining both strengths of the physical properties of each frequency band to achieve early and accurate classifications.

5.2.4 Importance of first acquisitions for early classification

Different classifications performed starting at various points in time have revealed interesting performance gaps. Indeed, the F1 score evolution plots illustrate how the model performs with the cumulative addition of data. Results indicate that for both sensors, removing acquisitions before mid-April does not significantly affect overall performance. This finding suggests that, in the context of this study, data gathered before mid-April is not crucial for early-season crop classification.

This insight is valuable for several reasons. Economically, reducing the need for early-season data acquisitions can lower costs, as radar products are not always freely available and can be expensive. Additionally, focusing on mid-April and later acquisitions allows

for more targeted and efficient use of resources, both in terms of data processing and storage.

5.3 Limitations and perspectives

Even though results look promising, several improvements could be applied to make global performance more accurate and extend the study on wider objectives.

Firstly, lacking of coherence computation possibilities for X-band data through TerraSAR-X consists of a huge loss of information. Coherence has been of great interest for the seedbed preparation detection and even though TerraSAR-X offers a longer revisit period showing fewer temporal details, a rough seedbed preparation period could have been estimated through coherence time series. Computing coherence on TerraSAR-X data could have been valuable to be able to compare time series and seedbed detection results, especially knowing that X-band is very sensitive to smaller scale roughness.

Even though seedbed preparation could be detected well for a few of the campaign fields, lack of larger validation for the complete Sentinel-1 footprint makes it hard to quantitatively estimate the accuracy of the detection, possibly highlighting some further limitations to coherence seedbed preparation detection.

As mentioned in Section 3.1.2, the months of March and April 2021 were drier than usual, allowing farmers to enter fields early in the season and sow at earlier stages. Additionally, the 2021 spring season was also one of the coldest recorded in Uccle since 1991, which might lead to slower or delayed vegetation growth on crop fields. These two factors emphasize the significant role of climate conditions in agriculture and can have a substantial impact on crop classification performance, raising concerns about the reproducibility of the model.

As observed with the combination of seedbed preparation period detection and classification results, the changes in soil roughness due to preparation for seeding can significantly improve classification accuracy (for example, beets could stand out early in the classification thanks to early seedbed preparation). An interesting future perspective could be to explore how classification accuracy evolves for each crop type when all crops are sown around the same period of the season.

Future studies could explore new horizons by evaluating performances of other classification algorithms and comparing them to Random Forest. It is also important to remember that this study only focuses on three very distinct culture types and that expanding the scope of seedbed preparation detection and classification to include other

crop types such as cereals, flax, etc., could have very different effects and provide new insights. Moreover, comparing the achieved results for 2021 with results that could be obtained for other years would help highlight potential inter-annual variations in weather or agricultural calendar for example.

Finally, this study could extend its scope to include the detection of tilled and untilled crop fields. This perspective is crucial for future agricultural management, particularly in the context of conservative agriculture, which is a focal point of the CAP.

CONCLUSION

The main objective of this master thesis was to evaluate the individual and combined contributions of X-band and C-band SAR data for early-season crop classification and detection of seedbed preparation activities with specific focus on the three most dominant spring crops in Wallonia: beet, maize, and potato. For this purpose, SAR acquisitions from TerraSAR-X and Sentinel-1 have been acquired and treated, backed up by an intensive field campaign conducted in 2021 in the area of Gembloux.

A seedbed preparation detection method based on VV coherence drop showed promising results in visualizing spatial and temporal distribution of seedbed preparation activities. Unfortunately, no clear validation data could reliably quantify the effectiveness of this method. Still, this first result provided valuable information for a better understanding of crop classification performances. Indeed, crop classification results using C-band or X-band have highlighted that the sowing period goes hand in hand with improvement of classification performance for that same period.

More globally, Random Forest classifications have mainly highlighted three important parameters in favor of better classifications: the need of an at least roughly balanced dataset, the importance of the sowing period and the non negligible role of high temporal resolution. Additionally, best overall results were achieved using a combination of TerraSAR-X and Sentinel-1 acquisitions, mixing small-scale roughness sensitivity of X-band data and the sensitivity to sub-surface soil layers of C-band data. In this particular case, beet fields could be accurately classified in early April, potato fields in late April and maize fields in the last weeks of May. This highlights the importance of the use of multi-frequency information to accurately classify selected crops as soon as possible in the growing season.

While several improvements can be applied to further evaluate the impact of certain choices made during this study (classification algorithm, minimum pixels thresholds, ...),

presented results offer a promising step forward in agricultural monitoring using SAR imagery. Seedbed detection and early-season crop classification can act as valuable tools for assisting farm advisory services, agricultural insurance companies, improving or build new agricultural policies, running yield or nitrogen uptake models, etc. All of these perspectives can play various roles in improving food security but also insuring a more sustainable agricultural sector.

BIBLIOGRAPHY

- Airbus. Radar Constellation (TerraSAR-X/PAZ) Processing Levels (Basic Image Products), n.d.
- AMS. Radar frequency band - Glossary of Meteorology, 2022. URL :https://glossary.ametsoc.org/wiki/Radar_frequency_band, (visited on 2023-11-25).
- G. Azzari, P. Grassini, J. I. R. Edreira, S. Conley, S. Mourtzinis, and D. B. Lobell. Satellite mapping of tillage practices in the North Central US region from 2005 to 2016. Remote Sensing of Environment, 221:417–429, 2019. doi: 10.1016/j.rse.2018.11.010.
- N. Baghdadi, S. Gaultier, and C. King. Retrieving surface roughness and soil moisture from synthetic aperture radar (SAR) data using neural networks. Canadian Journal of Remote Sensing, 28:701–711, 2002. doi: 10.5589/m02-066.
- N. Baghdadi, N. Holah, and M. Zribi. Soil moisture estimation using multi-incidence and multi-polarization ASAR data. International Journal of Remote Sensing, 27:1907–1920, 2006. doi: 10.1080/01431160500239032.
- N. Baghdadi, R. Cresson, M. El Hajj, R. Ludwig, and I. La Jeunesse. Estimation of soil parameters over bare agriculture areas from C-band polarimetric SAR data using neural networks. Hydrology and Earth System Sciences, 16:1607–1621, 2012a. doi: 10.5194/hess-16-1607-2012.
- N. Baghdadi, R. Cresson, E. Pottier, M. Aubert, M. Mehrez, A. Jacome, and S. Benabdallah. A Potential Use for the C-Band Polarimetric SAR Parameters to Characterize the Soil Surface Over Bare Agriculture Fields. IEEE Transactions on Geoscience and Remote Sensing, 50:3844–3858, 2012b. doi: 10.1109/tgrs.2012.2185934.
- H. Bahrami, H. McNairn, M. Mahdianpari, and S. Homayouni. A Meta-Analysis of Remote Sensing Technologies and Methodologies for Crop Characterization. Remote Sensing, 14:5633, 2022. doi: 10.3390/rs14225633.
- BCGMS. Bulletin agrométéorologique (Belgique) - Situation au 1er juillet, 2021.
- BCGMS. Bulletin agrométéorologique (Belgique) - Situation au 15 mai, 2024.
- I. Becker-Reshef, E. Vermote, M. Lindeman, and C. Justice. A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. Remote Sensing of Environment, 114:1312–1323, 2010. doi: 10.1016/j.rse.2010.01.010.

- P. Bertin. LBRAI2106 - Phytotechnie, Faculté des bioingénieurs, UCLouvain. 2023.
- G. Biau and E. Scornet. A random forest guided tour. TEST, 25:264–268, 2016. doi: 10.1007/s11749-016-0488-0.
- C. Bielders, F. Colard, H. Cordonnier, A. Degré, M.-F. Destain, N. Feltz, and A. Maignard. Convention de recherche d’intérêt général Gestion Intégrée Sol Erosion Ruissellement : Rapport Final. 2011.
- E. Y. Boateng, J. Otoo, and D. A. Abaye. Basic Tenets of Classification Algorithms K-Nearest-Neighbor, Support Vector Machine, Random Forest and Neural Network: A Review. Journal of Data Analysis and Information Processing, 08:341–357, 2020. doi: 10.4236/jdaip.2020.84020.
- L. Breiman. Random Forests. Machine Learning, 45:5–32, 2001. doi: 10.1023/a:1010933404324.
- J. B. Campbell and R. H. Wynne. Introduction to Remote Sensing. Guilford Publications, Incorporated, 5 edition, 2011.
- N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 16:321–357, 2002. doi: 10.1613/jair.953.
- M. Choker, N. Baghdadi, M. Zribi, M. El Hajj, S. Paloscia, N. Verhoest, H. Lievens, and F. Mattia. Evaluation of the Oh, Dubois and IEM Backscatter Models Using a Large Dataset of SAR Data and Experimental Soil Measurements. Water, 9:38–38, 2017. doi: 10.3390/w9010038.
- G. Chomé, P. Baret, and P. Defourny. Mapping farming practices in Belgian intensive cropping systems from Sentinel-1 SAR time-series. 2016.
- C. A. Clark. The important role of soil texture on water, 2023. <https://cropsandsoils.extension.wisc.edu/articles/the-important-role-of-soil-texture-on-water/>, (visited on 2024-03-05).
- C. Cortes and V. Vapnik. Support-vector networks. Machine Learning, 20:273–297, 1995. doi: 10.1007/bf00994018.
- T. Cover and P. Hart. Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13:21–27, 1967. doi: 10.1109/tit.1967.1053964.
- CRA-W. AGROMET : a web-based weather platform for agricultural decision support systems, n.d. <https://www.cra.wallonie.be/en/agromet>, (visited on 2024-03-22).
- M. De Vroey, J. Radoux, and P. Defourny. Grassland Mowing Detection Using Sentinel-1 Time Series: Potential and Limitations. Remote Sensing, 13:348, 2021. doi: 10.3390/rs13030348.
- P. Defourny. LBIRE2102 - Géomatique appliquée, Faculté des bioingénieurs, UCLouvain. 2022.
- Diplomatie Belgium. Agriculture and fisheries, 2023. URL :<https://diplomatie.belgium.be/en/policy/coordination-european-affairs/policy-belgium-within-eu/agriculture-and-fisheries>, (visited on 2023-11-25).
- DLR. TerraSAR-X: Basic Product Specification Document, n.d.
- EO College. Basic Principles of Radar Backscatter. 2023. <https://eo-college.org/courses/basic-principles-of-radar-backscatter/>, (visited on 2023-10-15).

- ESA. SNAP – ESA Sentinel Application Platform v9.0.0, 2022. <https://step.esa.int/main/download/snap-download/>, (visited on 2022-03-24).
- ESA. Sentinel-1 - Instrument Payload - Sentinel Online - Sentinel Online, n.d.a. <https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-1/instrument-payload>, (visited on 2024-03-20).
- ESA. TerraSAR-X and TanDEM-X - Earth Online, n.d.b. <https://earth.esa.int/eogateway/missions/terrasar-x-and-tandem-x>, (visited on 2024-03-20).
- European Comission. Key policy objectives of the CAP 2023-27, n.d. URL :https://agriculture.ec.europa.eu/common-agricultural-policy/cap-overview/cap-2023-27/key-policy-objectives-cap-2023-27_en, (visited on 2023-11-26).
- European Court of Auditors. The Land Parcel Identification System: a useful tool to determine the eligibility of agricultural land – but its management could be further improved, 2016.
- FAO. The State of Food Security and Nutrition in the World 2023. Food And Agriculture Organization Of The United Nations, 2023.
- J. A. Foley, N. Ramankutty, K. A. Brauman, E. S. Cassidy, J. S. Gerber, M. Johnston, N. D. Mueller, C. O’Connell, D. K. Ray, P. C. West, C. Balzer, E. M. Bennett, S. R. Carpenter, J. Hill, C. Monfreda, S. Polasky, J. Rockström, J. Sheehan, S. Siebert, D. Tilman, and D. P. M. Zaks. Solutions for a cultivated planet. Nature, 478:337–342, 2011. doi: 10.1038/nature10452.
- A. K. Fung, Z. B. Li, and K. F. Chen. Backscattering from a randomly rough dielectric surface. IEEE Transactions on Geoscience and Remote Sensing, 30:356–369, 1992. doi: 10.1109/36.134085.
- Geoportail Wallonie. Géoportail de la Wallonie, n.d. <https://geoportail.wallonie.be/home.html>, (visited on 2024-03-20).
- D. Gunay. Random Forest, 2023. <https://medium.com/@denizgunay/random-forest-af5bde5d7e1e>, (visited on 2024-04-11).
- R. Hadria, B. Duchemin, F. Baup, T. Le Toan, A. Bouvet, G. Dedieu, and M. Le Page. Combined use of optical and radar satellite data for the detection of tillage and irrigation operations: Case study in Central Morocco. Agricultural Water Management, 96:1120–1127, 2009. doi: 10.1016/j.agwat.2009.02.010.
- Harikrishnan. Confusion Matrix, Accuracy, Precision, Recall, F1 Score, 2019. <https://medium.com/analytics-vidhya/confusion-matrix-accuracy-precision-recall-f1-score-ade299cf63cd>, (visited on 2024-04-12).
- Haviluddin, M. Iqbal, G. Mahendra Putra, N. Puspitasari, H. Jati Setyadi, F. Andika Dwiyanto, A. Prasetya Wibawa, and R. Alfred. A Performance Comparison of Euclidean, Manhattan and Minkowski Distances in K-Means Clustering. 2020. doi: 10.1109/icsitech49800.2020.9392053.
- K. He, Q. Yang, X. Shen, and E. N. Anagnostou. Brief communication: Western Europe flood in 2021 – mapping agriculture flood exposure from synthetic aperture radar (SAR). Natural Hazards and Earth System Sciences, 22:2921–2927, 2022. doi: 10.5194/nhess-22-2921-2022.
- P. Huilgol. Accuracy vs. f1-score, 2019. <https://medium.com/analytics-vidhya/accuracy-vs-f1-score-6258237beca2>, (visited on 2024-04-23).

- IRM. Bilans climatologiques de 2021 - Printemps, 2021. <https://www.meteo.be/fr/climat/climat-de-la-belgique/bilans-climatologiques/2021/printemps>, (visited on 2024-03-06).
- T. Jayne and S. Rashid. The Value of Accurate Crop Production Forecasts. 2010.
- S. Khabbazan, P. Vermunt, S. Steele-Dunne, L. Ratering Arntz, C. Marinetti, D. van der Valk, L. Iannini, R. Molijn, K. Westerdijk, and C. van der Sande. Crop Monitoring Using Sentinel-1 Data: A Case Study from The Netherlands. Remote Sensing, 11:1887, 2019. doi: 10.3390/rs11161887.
- G. Lemoine, H. de Groof, H. van Leeuwen, and J. Bakker. Monitoring agricultural land preparation activities with ERS-1 SAR data. IEEE International Symposium on Geoscience and Remote Sensing (IGARSS), 2002. doi: 10.1109/igarss.1995.521188.
- G. Louppe, L. Wehenkel, A. Sutera, and P. Geurts. Understanding variable importances in forests of randomized trees. Advances in Neural Information Processing Systems, 26, 2013.
- M. Luotamo, M. Yli-Heikkilä, and A. Klami. Density Estimates as Representations of Agricultural Fields for Remote Sensing-Based Monitoring of Tillage and Vegetation Cover. Applied Sciences, 12:679, 2022. doi: 10.3390/app12020679.
- D. Mandal, V. Kumar, D. Ratha, S. Dey, A. Bhattacharya, J. M. Lopez-Sanchez, H. McNairn, and Y. S. Rao. Dual polarimetric radar vegetation index for crop growth monitoring using sentinel-1 SAR data. Remote Sensing of Environment, 247:111954, 2020. doi: 10.1016/j.rse.2020.111954.
- Mathworks. Support Vector Machines (SVM), n.d. <https://ch.mathworks.com/fr/discovery/support-vector-machine.html>, (visited on 2024-04-11).
- F. Mattia, G. Satalino, V. R. N. Pauwels, and A. Loew. Soil moisture retrieval through a merging of multi-temporal L-band SAR data and hydrologic modelling. Hydrology and Earth System Sciences, 13:343–356, 2009. doi: 10.5194/hess-13-343-2009.
- H. McNairn, J. B. Boisvert, D. J. Major, G. , R. B. Brown, and A. M. Smith. Identification of Agricultural Tillage Practices from C-Band Radar Backscatter. Canadian Journal of Remote Sensing, 22:154–162, 1996. doi: 10.1080/07038992.1996.10874649.
- H. McNairn, D. Wood, Q. Gwyn, R. Brown, and F. Charbonneau. Mapping Tillage and Crop Residue Management Practices with RADARSAT. Canadian Journal of Remote Sensing, 24:28–35, 1998. doi: 10.1080/07038992.1998.10874688.
- H. McNairn, C. Champagne, J. Shang, D. Holmstrom, and G. Reichert. Integration of optical and Synthetic Aperture Radar (SAR) imagery for delivering operational annual crop inventories. ISPRS Journal of Photogrammetry and Remote Sensing, 64:434–449, 2009. doi: 10.1016/j.isprsjprs.2008.07.006.
- H. McNairn, A. Kross, D. Lapen, R. Caves, and J. Shang. Early season monitoring of corn and soybeans with TerraSAR-X and RADARSAT-2. International Journal of Applied Earth Observation and Geoinformation, 28:252–259, 2014. doi: 10.1016/j.jag.2013.12.015.
- H. McNairn, L. D. Robertson, M. , S. Ihuoma, X. Jiao, and P. Joosse. Coherent Change Detection to Monitor Tillage. IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium, 2022. doi: 10.1109/igarss46834.2022.9883744.
- A. Mucherino, P. J. Papajorgji, and P. M. Pardalos. k-Nearest Neighbor Classification. Data Mining in Agriculture, pages 83–106, 2009. doi: 10.1007/978-0-387-88615-2_4.

- NASA. SAR Data Pre-Processing Steps, n.d. https://www.earthdata.nasa.gov/s3fs-public/imported/pre-processing_onepager_ROUGH.pdf, (visited on 2024-04-15).
- R. Nasirzadehdizaji, Z. Cakir, F. Balik Sanli, S. Abdikan, A. Pepe, and F. Calò. Sentinel-1 interferometric coherence and backscattering analysis for crop monitoring. Computers and Electronics in Agriculture, 185:106118, 2021. doi: 10.1016/j.compag.2021.106118.
- S. Paloscia, P. Pampaloni, S. Pettinato, and E. Santi. A Comparison of Algorithms for Retrieving Soil Moisture from ENVISAT/ASAR Images. IEEE Transactions on Geoscience and Remote Sensing, 46: 3274–3284, 2008. doi: 10.1109/tgrs.2008.920370.
- M. Preiss and N. Stacy. Coherent Change Detection: Theoretical Description and Experimental Results. 2006.
- L. D. Robertson, H. McNairn, M. , X. Jiao, S. Ihuoma, and P. Joosse. Monitoring autumn agriculture activities using Synthetic Aperture Radar (SAR) and coherence change detection. Heliyon, 9:e17322–e17322, 2023. doi: 10.1016/j.heliyon.2023.e17322.
- S. M. Sabery, A. Bystrov, M. Navarro-Cía, P. Gardner, and M. Gashinova. Study of Low Terahertz Radar Signal Backscattering for Surface Identification. Sensors, 21:2954–2954, 2021. doi: 10.3390/s21092954.
- G. Satalino, F. Mattia, M. Davidson, T. L. Toan, G. Pasquariello, and M. Borgeaud. On current limits of soil moisture retrieval from ERS-SAR data. IEEE Transactions on Geoscience and Remote Sensing, 40:2438–2447, 2002. doi: 10.1109/tgrs.2002.803790.
- G. Satalino, F. Mattia, A. Balenzano, F. P. Lovergine, M. Rinaldi, A. Pio, S. Ruggieri, D. Alfonso, V. P. Gómez, Éric Ceschia, M. Planells, T. L. Toan, A. Ruiz, and J. Moreno. Sentinel-1 Sentinel-2 Data for Soil Tillage Change Detection. IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, 2018.
- J. Shang, J. Liu, V. Poncos, X. Geng, B. Qian, Q. Chen, T. Dong, D. Macdonald, T. Martin, J. Kovacs, and D. Walters. Detection of Crop Seeding and Harvest through Analysis of Time-Series Sentinel-1 Interferometric SAR Data. Remote Sensing, 12:1551, 2020. doi: 10.3390/rs12101551.
- SPW. Carte des Principaux Types de Sols de Wallonie à 1/250000, n.d. <https://geoportail.wallonie.be/catalogue/64bbc088-367c-485c-bd7c-d2d08baedf9d.html>, (visited on 2024-03-20).
- SPW Agriculture. Régions agricoles de Wallonie, n.d. URL :https://etat-agriculture.wallonie.be/contents/indicatorsheets/EAW-A_I_d_2.html, (visited on 2024-03-05).
- H. Srivastava, P. Patel, M. Manchanda, and S. Adiga. Use of multiincidence angle RADARSAT-1 SAR data to incorporate the effect of surface roughness in soil moisture estimation. IEEE Transactions on Geoscience and Remote Sensing, 41:1638–1640, 2003. doi: 10.1109/tgrs.2003.813356.
- Statbel. Chiffres clés de l’agriculture 2021. 2021. URL : <https://agriculture.wallonie.be/documents/20182/43924/Chiffres-clés-de-l-agriculture-2021--L-agriculture-belge-en-chiffres.pdf/903f8b2d-954d-426f-b93d-3da5305f5e22>, (visited on 2023-11-25).
- Statbel. Key Figures on Agriculture 2023 | Statbel, 2023. URL :<https://statbel.fgov.be/en/news/key-figures-agriculture-2023#:~:text=The%20utilised%20agricultural%20area%20accounts>, (visited on 2023-11-25).
- United Nations. Sustainable Development Goals, n.d. <https://sdgs.un.org/goals>, (visited on 2024-06-08).

- K. Van Tricht, A. Gobin, S. Gilliams, and I. Piccard. Synergistic Use of Radar Sentinel-1 and Optical Sentinel-2 Imagery for Crop Mapping: A Case Study for Belgium. Remote Sensing, 10:1642, 2018. doi: 10.3390/rs10101642.
- K. Voormansik, K. Zalite, I. Sünter, T. Tamm, K. Koppel, T. Verro, A. Brauns, D. Jakovels, and J. Praks. Separability of Mowing and Ploughing Events on Short Temporal Baseline Sentinel-1 Coherence Time Series. Remote Sensing, 12:3784, 2020. doi: 10.3390/rs12223784.
- M. Vreugdenhil, C. Navacchi, B. Bauer-Marschallinger, S. Hahn, S. Steele-Dunne, I. Pfeil, W. Dorigo, and W. Wagner. Sentinel-1 Cross Ratio and Vegetation Optical Depth: A Comparison over Europe. Remote Sensing, 12:3404, 2020. doi: 10.3390/rs12203404. URL <https://www.mdpi.com/2072-4292/12/20/3404>.
- J. Warren and R. Taylor. Managing Soil Compaction - Oklahoma State University, 03 2017. <https://extension.okstate.edu/fact-sheets/managing-soil-compaction.html>, (visited on 2024-03-05).
- Y. Zhang. Support Vector Machine Classification Algorithm and Its Application. Communications in Computer and Information Science, 308:179–186, 2012. doi: 10.1007/978-3-642-34041-3_27.
- J. Álvarez-Mozos, J. Casali, M. González-Audicana, and N. Verhoest. Assessment of the operational applicability of RADARSAT-1 data for surface soil moisture estimation. IEEE Transactions on Geoscience and Remote Sensing, 44:913–924, 2006. doi: 10.1109/tgrs.2005.862248.

A Materials and methods

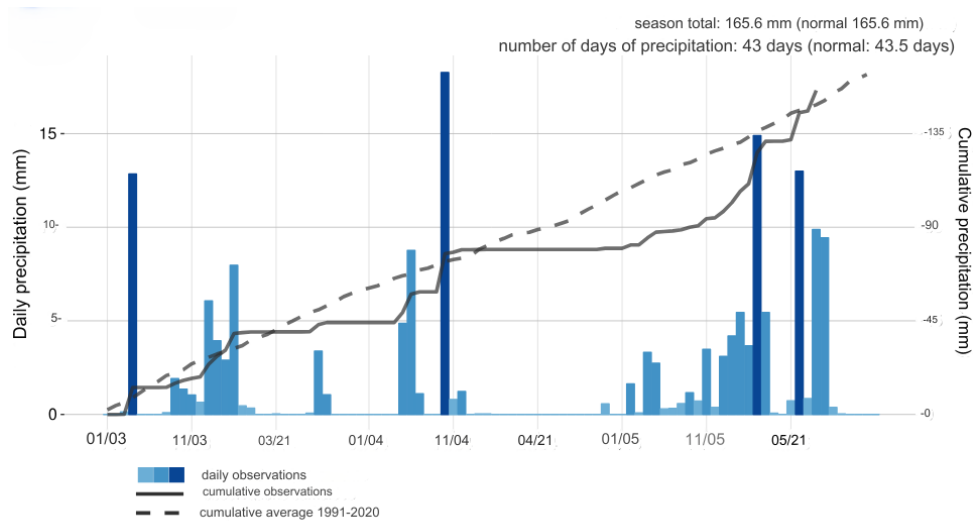


Figure 8.1: Daily and cumulative precipitation in Uccle for the 2021 spring season. Blue bars are representing daily observations with darkness depending on precipitation intensity. The dashed line shows the cumulative average between 1991 and 2020 while the full line shows cumulative observations (IRM, 2021).

B Results

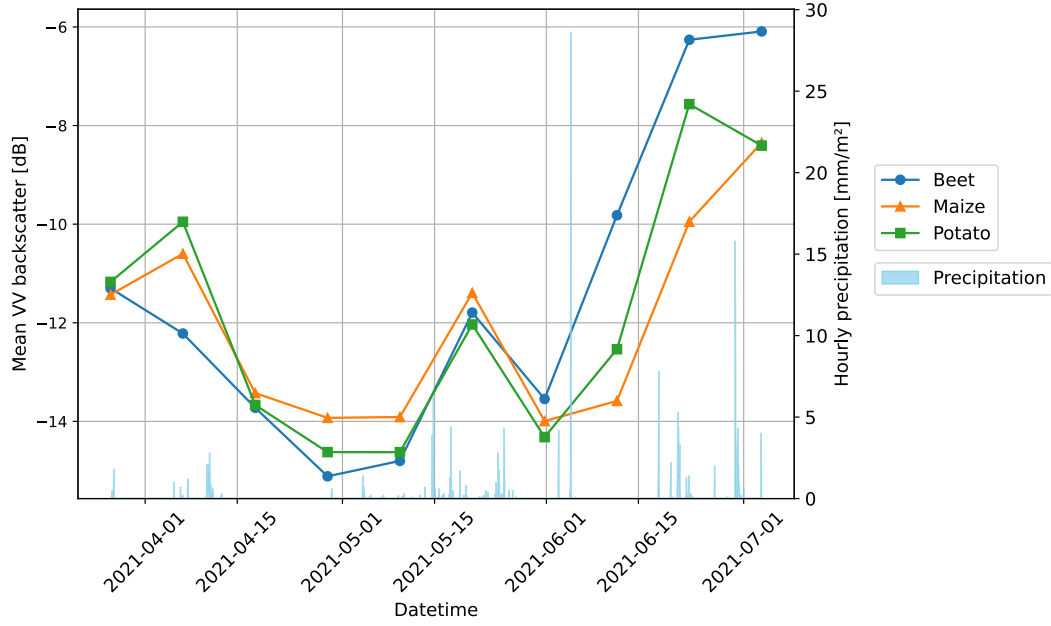


Figure 8.2: Mean evolution of the TerraSAR-X VV backscatter (σ_{VV}^0) and hourly precipitation through the 2021 spring season for all parcels of each studied culture (beet, maize, potato).

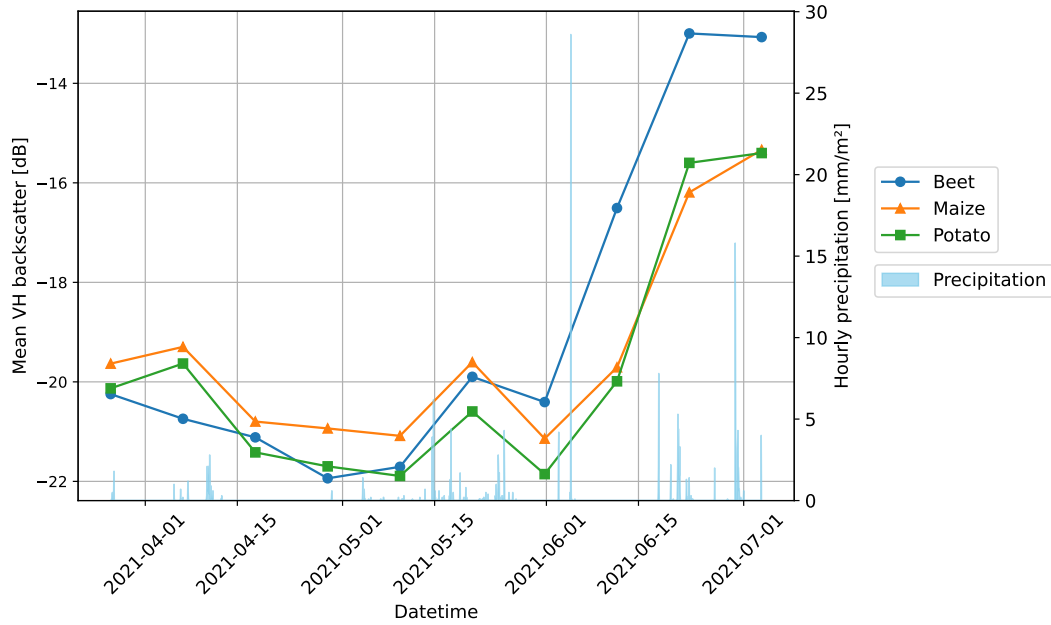


Figure 8.3: Mean evolution of the TerraSAR-X VH backscatter (σ_{VH}^0) and hourly precipitation through the 2021 spring season for all parcels of each studied culture (beet, maize, potato).

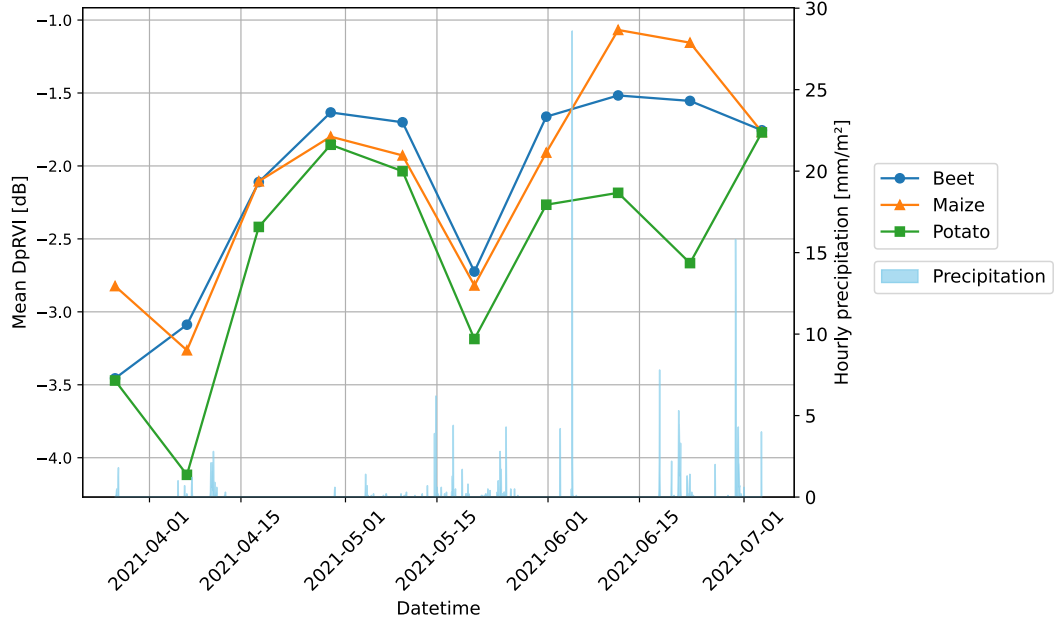


Figure 8.4: Mean evolution of the TerraSAR-X DpRVI and hourly precipitation through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).

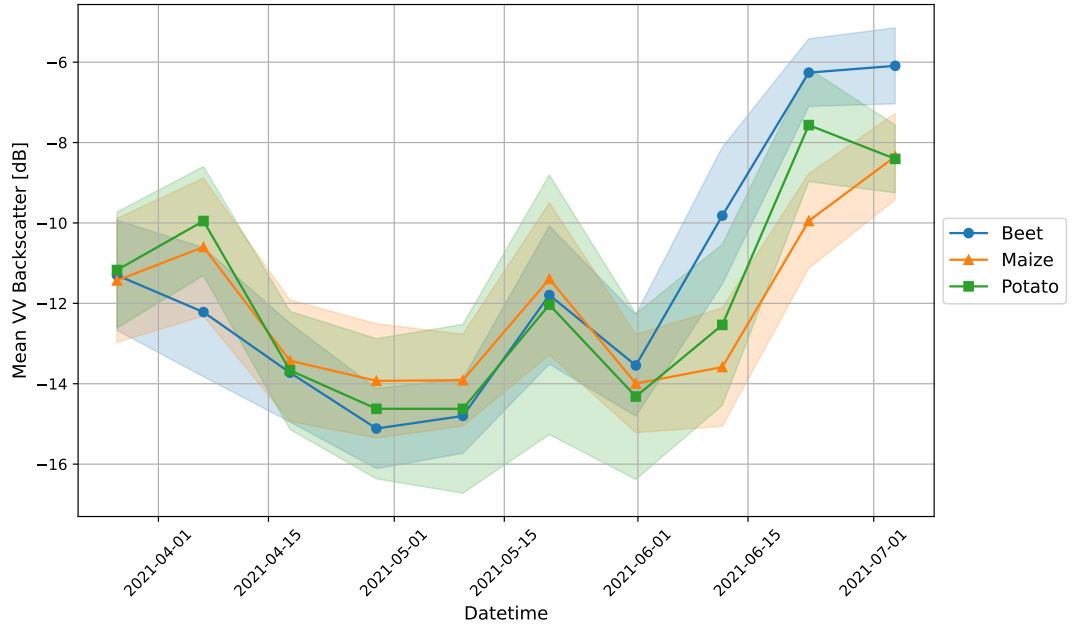


Figure 8.5: Mean evolution of the TerraSAR-X VV backscatter (σ_{VV}^0) with standard deviation buffer through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).

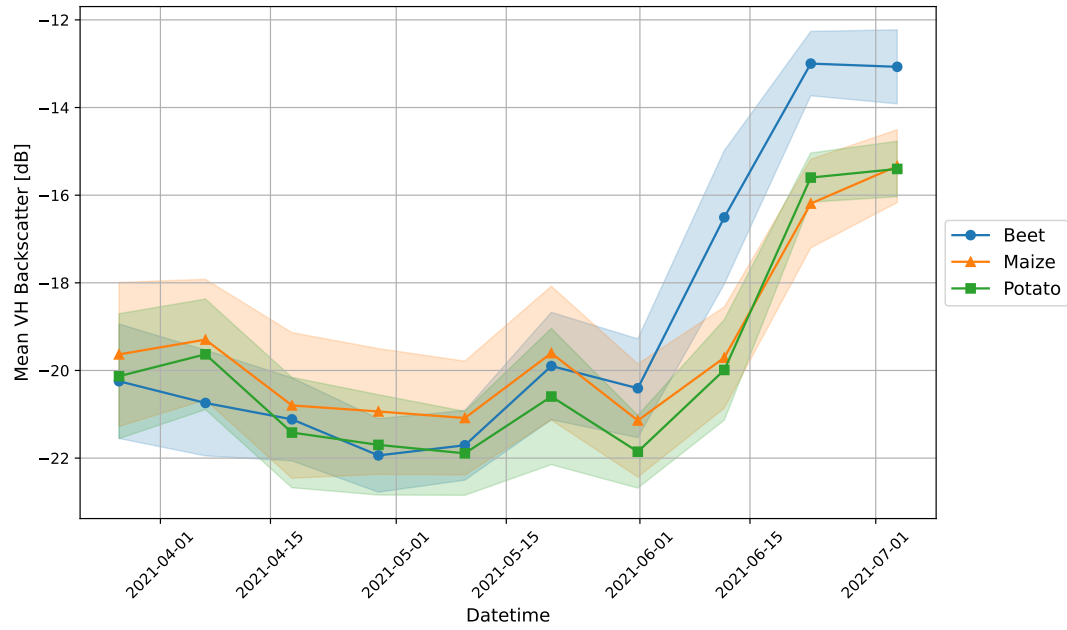


Figure 8.6: Mean evolution of the TerraSAR-X VH backscatter (σ_{VH}^0) with standard deviation buffer through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).

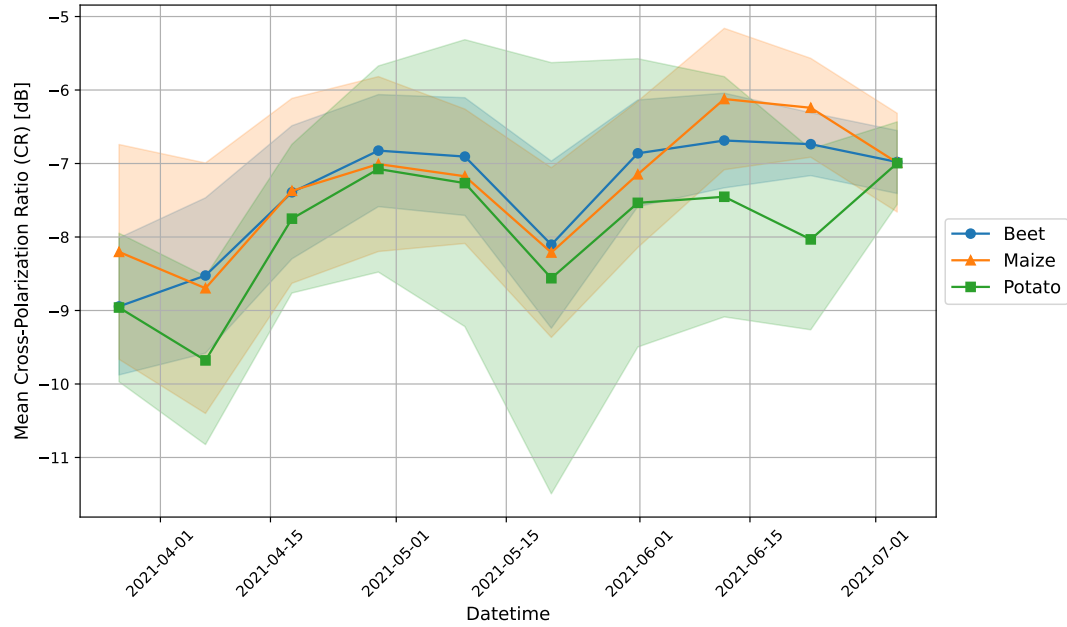


Figure 8.7: Mean evolution of the TerraSAR-X CR with standard deviation buffer through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).

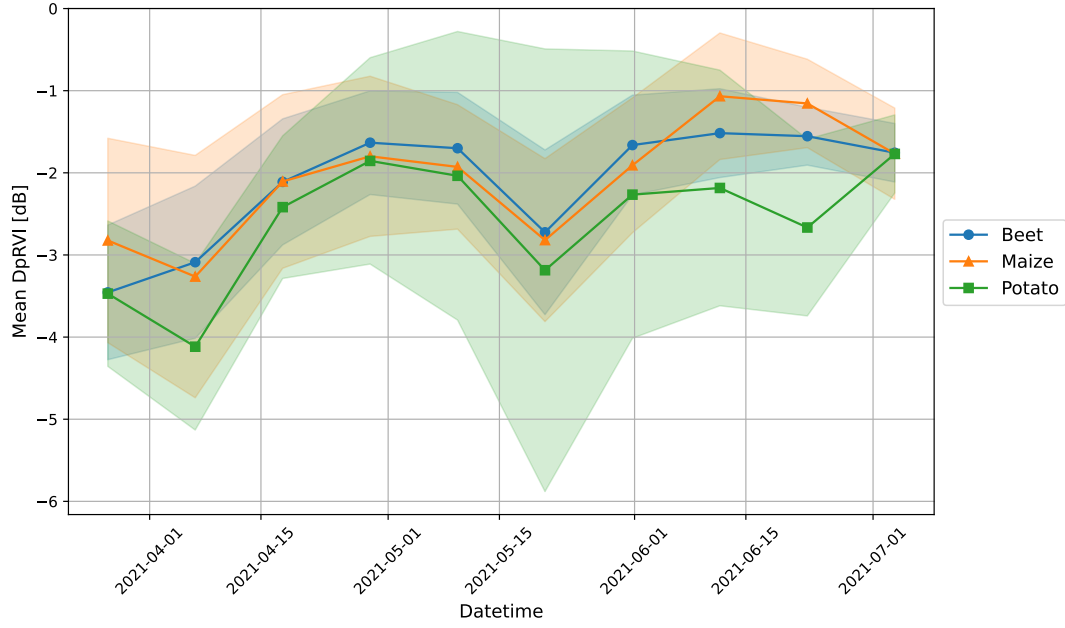


Figure 8.8: Mean evolution of the TerraSAR-X DpRVI with standard deviation buffer through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).

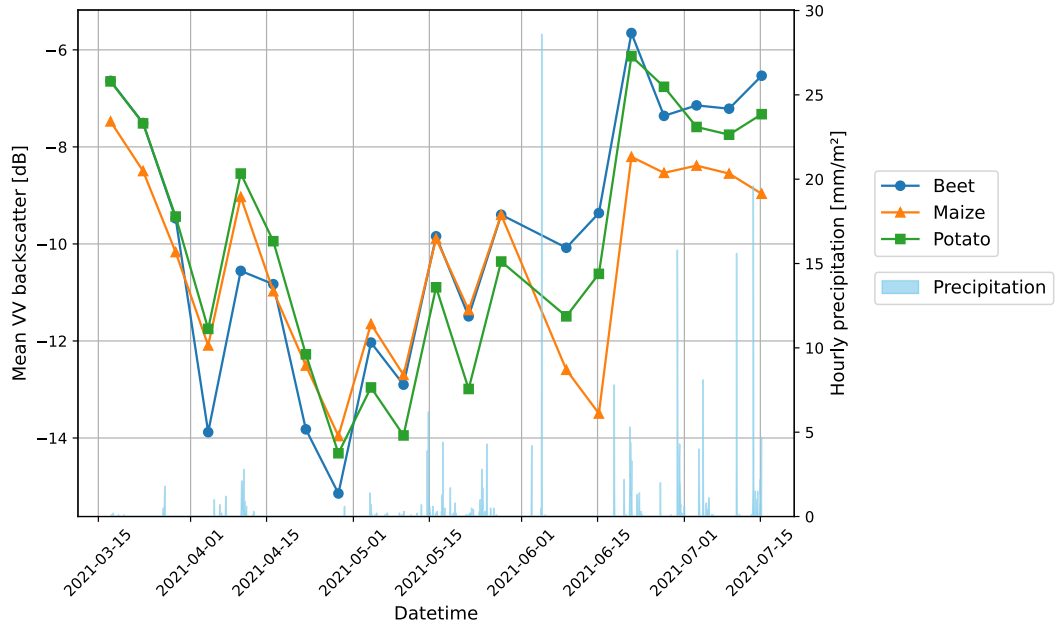


Figure 8.9: Mean evolution of the Sentinel-1 VV backscatter (σ_{VV}^0) and hourly precipitation through the 2021 spring season for all parcels of each studied culture (beet, maize, potato)).

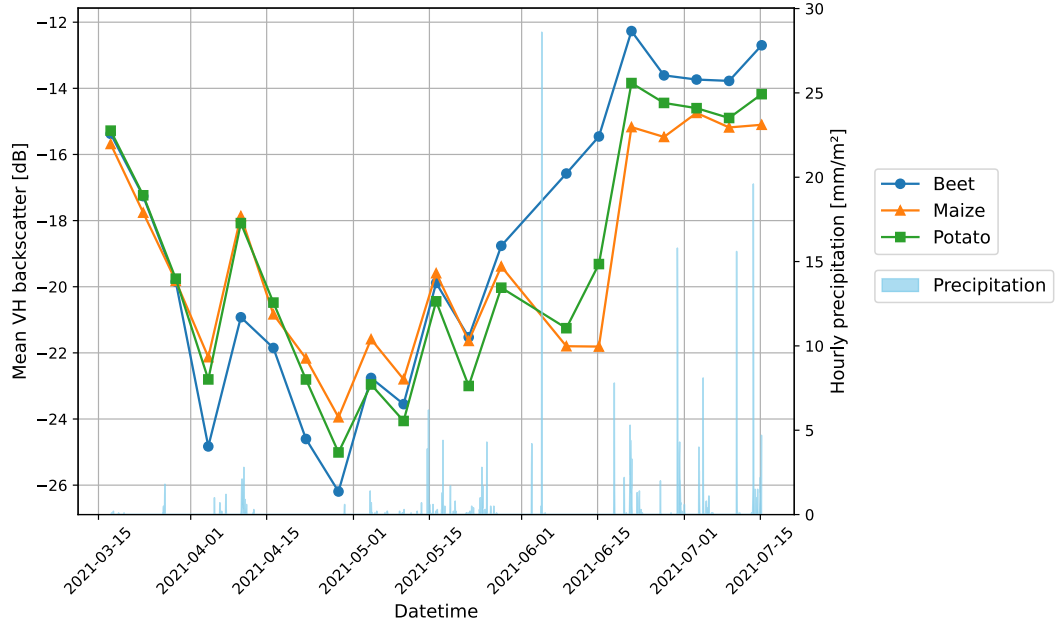


Figure 8.10: Mean evolution of the Sentinel-1 VH backscatter (σ_{VH}^0) and hourly precipitation through the 2021 spring season for all parcels of each studied culture (beet, maize, potato).

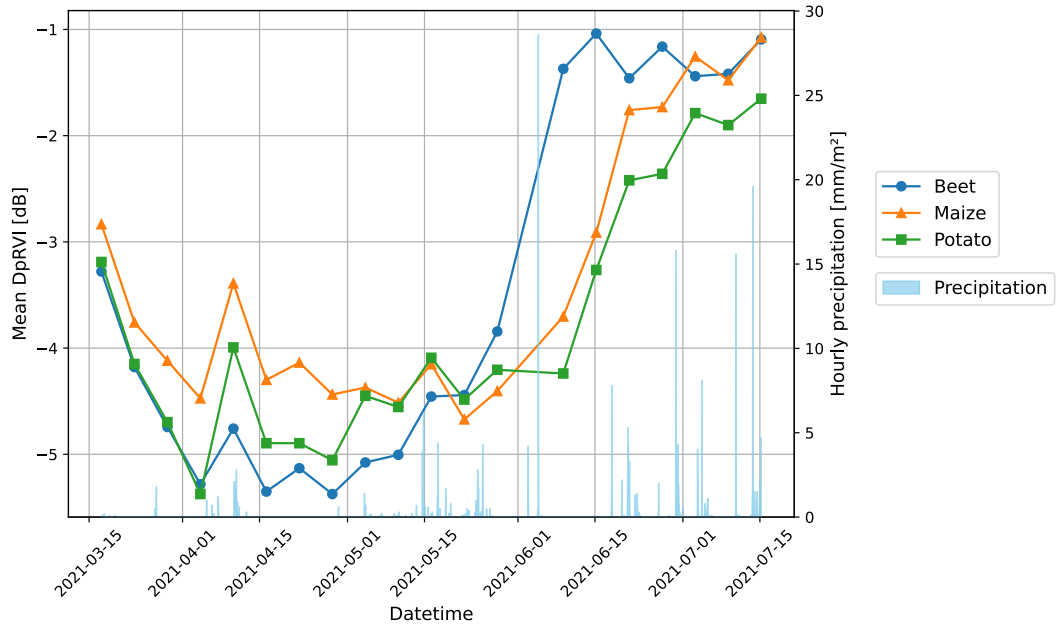


Figure 8.11: Mean evolution of the Sentinel-1 DpRVI and hourly precipitation through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).

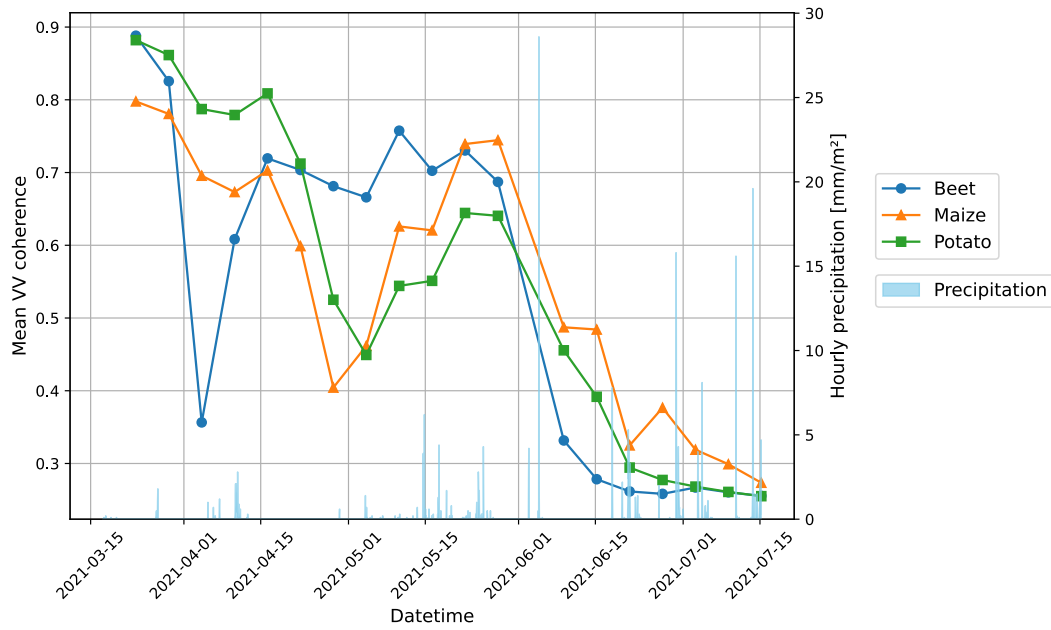


Figure 8.12: Mean evolution of the Sentinel-1 VV coherence and hourly precipitation through the 2021 spring season for all parcels of each studied culture (beet, maize, potato).

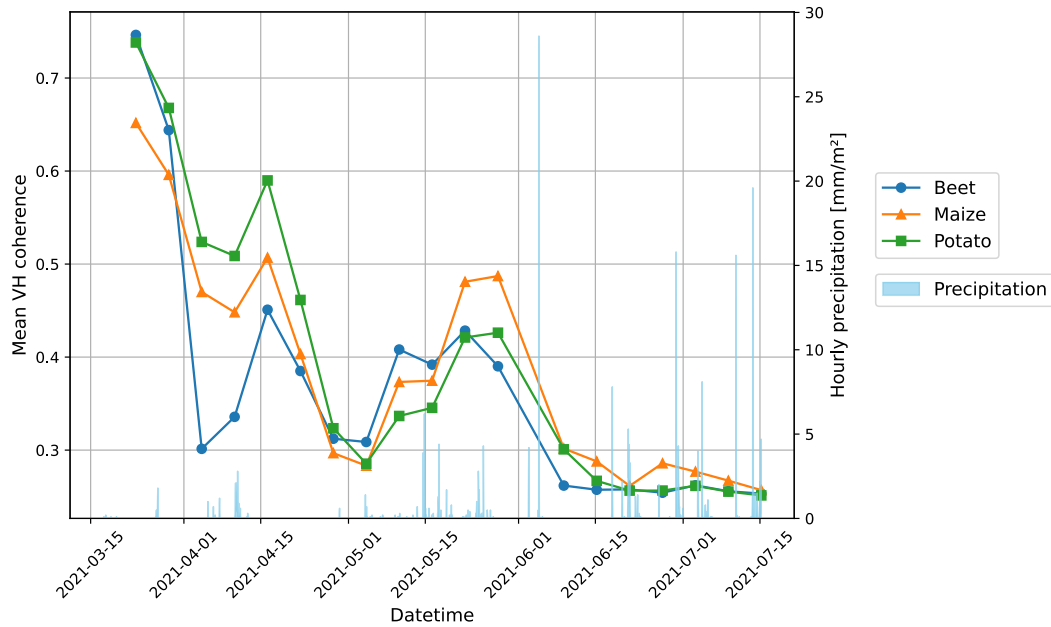


Figure 8.13: Mean evolution of the Sentinel-1 VH coherence and hourly precipitation through the 2021 spring season for all parcels of each studied culture (beet, maize, potato).

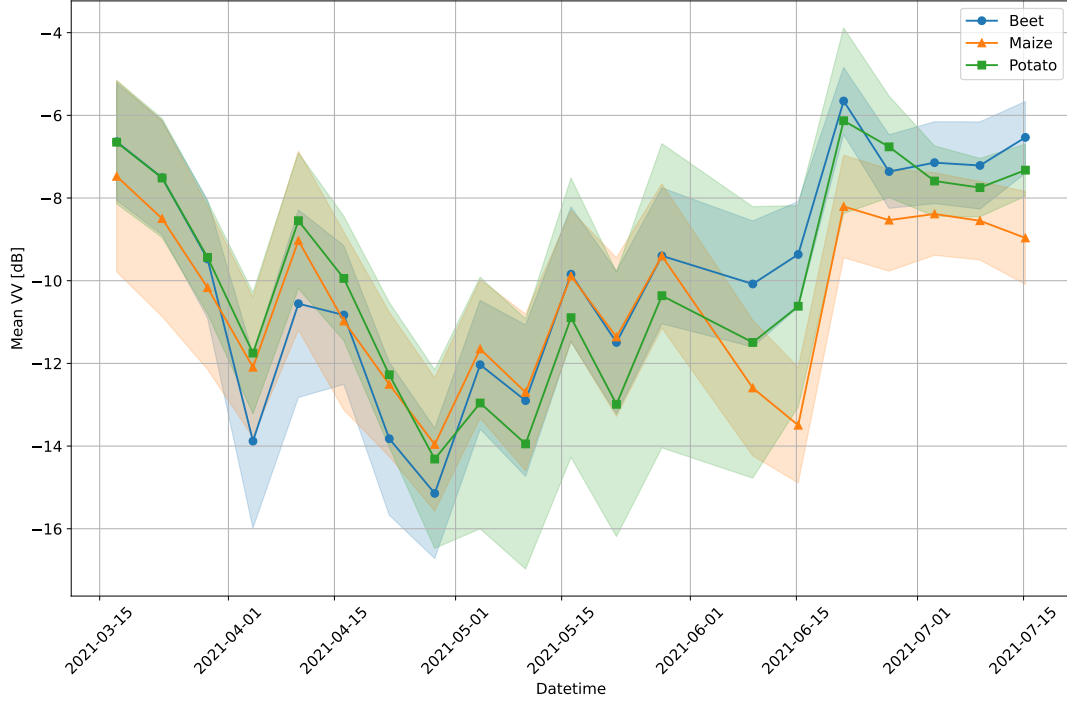


Figure 8.14: Mean evolution of the Sentinel-1 VV backscatter (σ_{VV}^0) with standard deviation buffer through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).

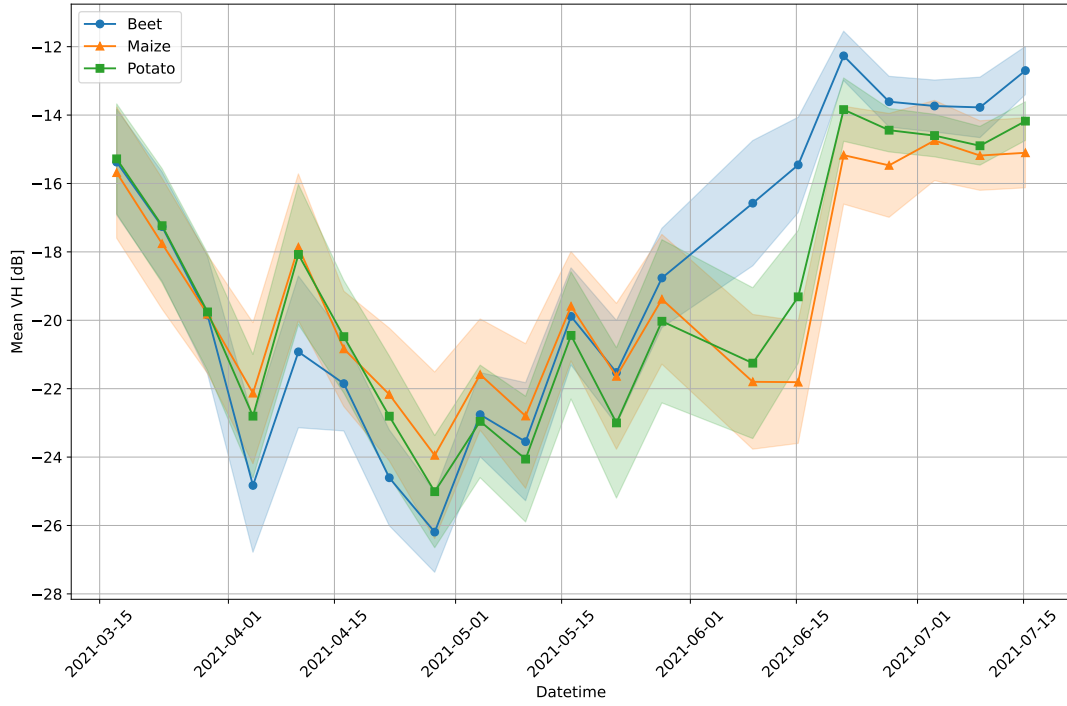


Figure 8.15: Mean evolution of the Sentinel-1 VH backscatter (σ_{VH}^0) with standard deviation buffer through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).

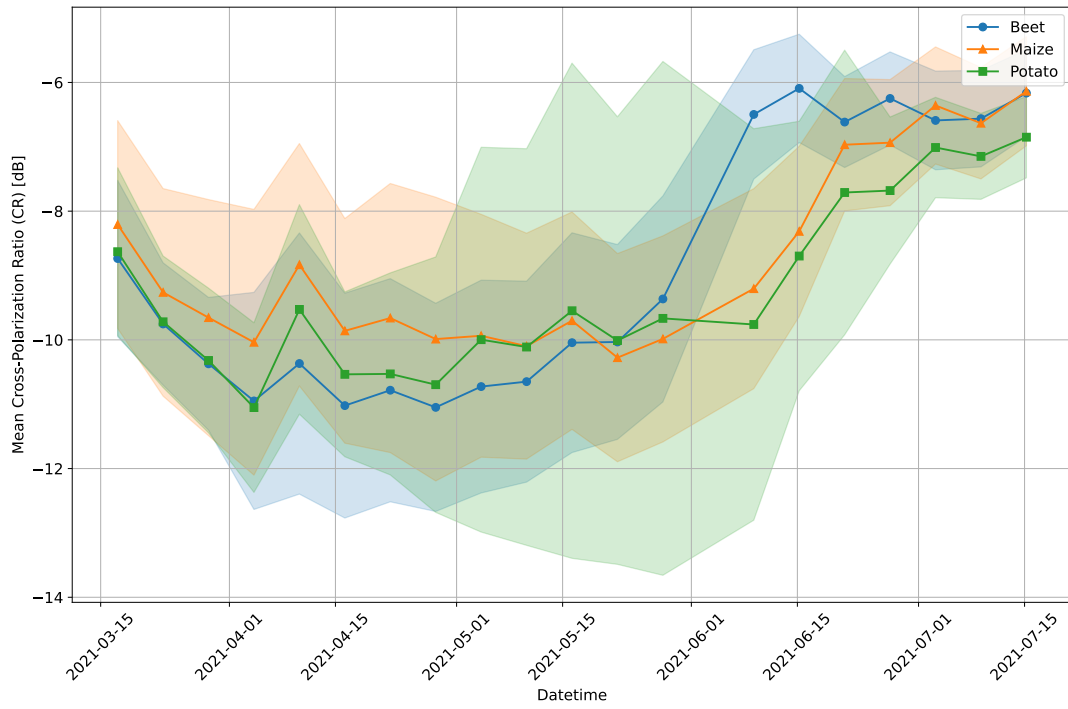


Figure 8.16: Mean evolution of the Sentinel-1 CR with standard deviation buffer through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).

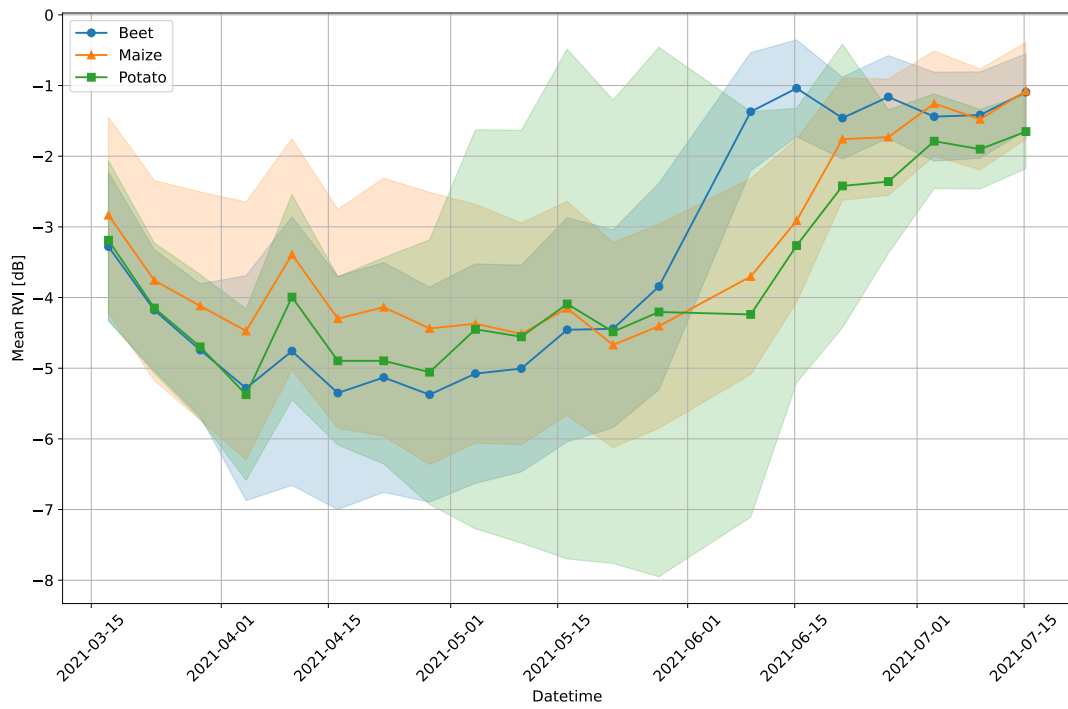


Figure 8.17: Mean evolution of the Sentinel-1 DpRVI with standard deviation buffer through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).

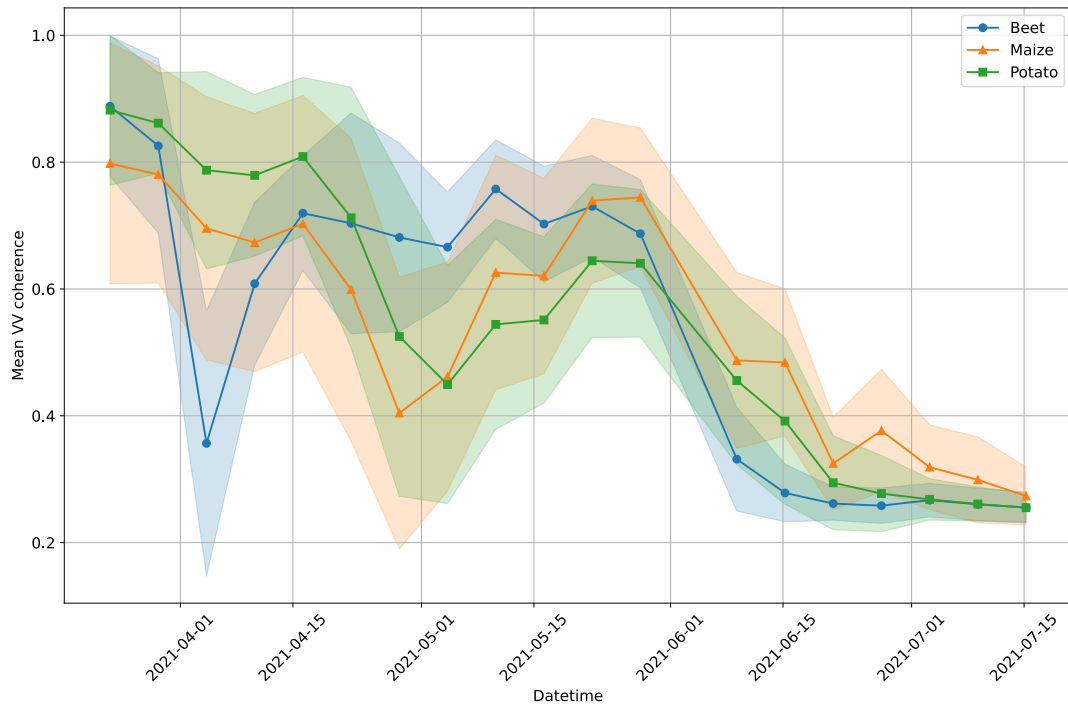


Figure 8.18: Mean evolution of the Sentinel-1 VV coherence with standard deviation buffer through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).

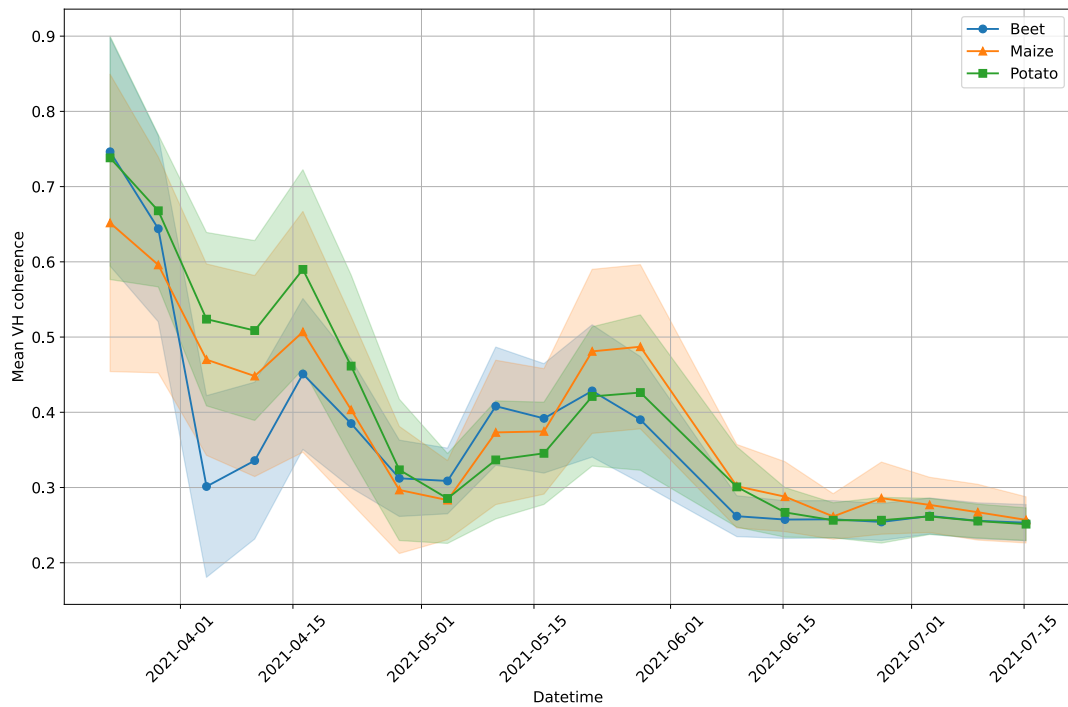


Figure 8.19: Mean evolution of the Sentinel-1 VH coherence with standard deviation buffer through the 2021 spring season for all parcels of each studied culture (beet, potato, maize).

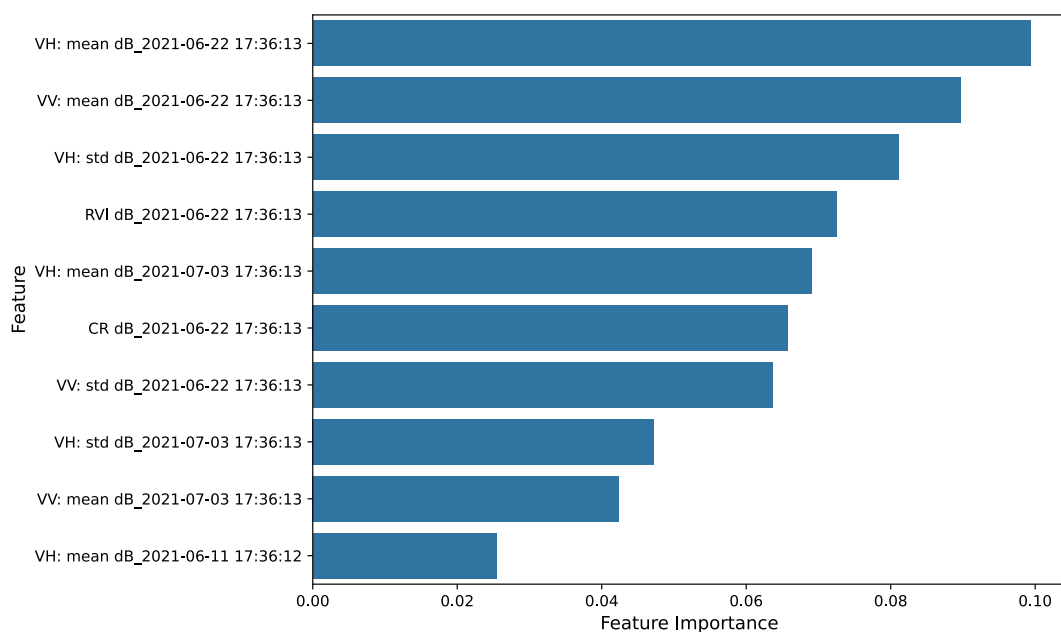


Figure 8.20: Top ten of most important features for the Random Forest classification of beet, maize, and potato crops using TerraSAR-X data on the area of interest, while maintaining a minimum threshold of 100 pixels of image per field.

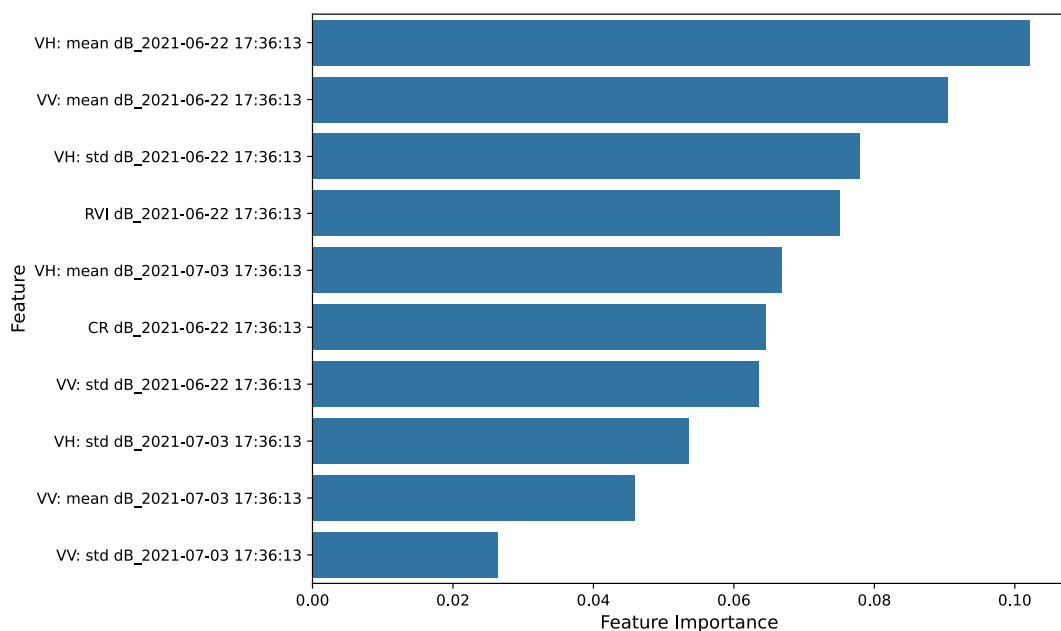


Figure 8.21: Top ten of most important features for the Random Forest classification of beet, maize, and potato crops using TerraSAR-X data on the area of interest, while maintaining a minimum threshold of 150 pixels of image per field.

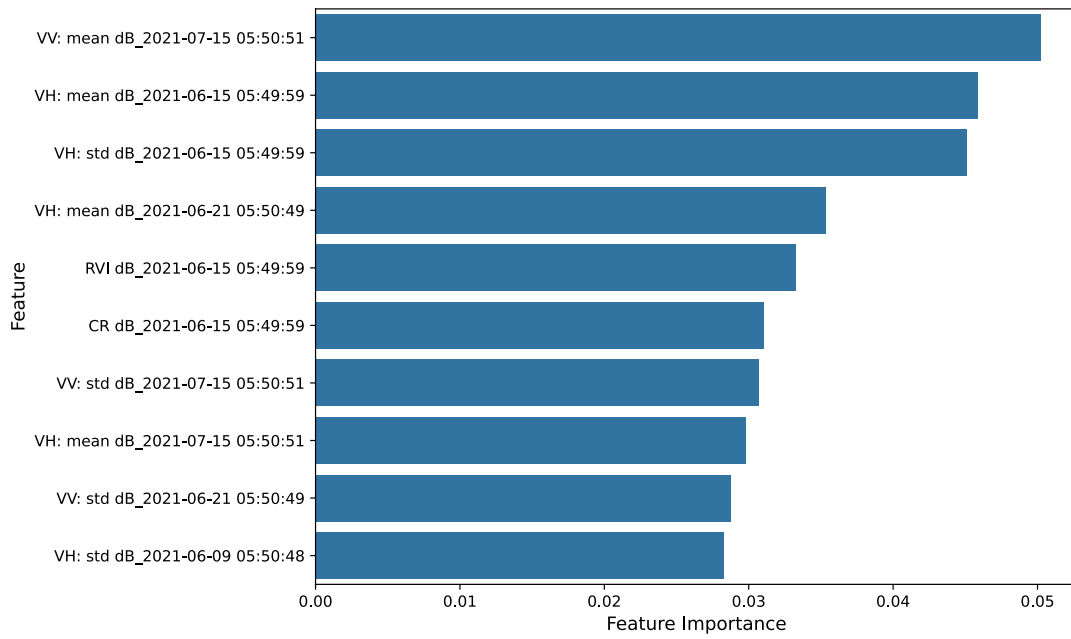


Figure 8.22: Top ten of most important features for the Random Forest classification of beet, maize, and potato crops using Sentinel-1 data on the area of interest, while maintaining a minimum threshold of 100 pixels of image per field.

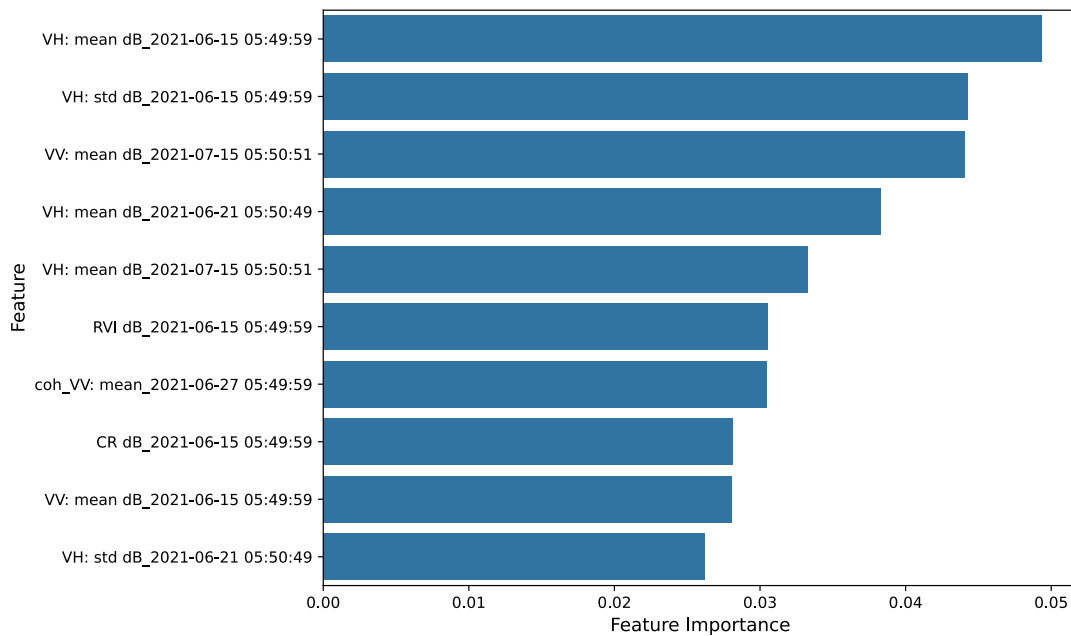
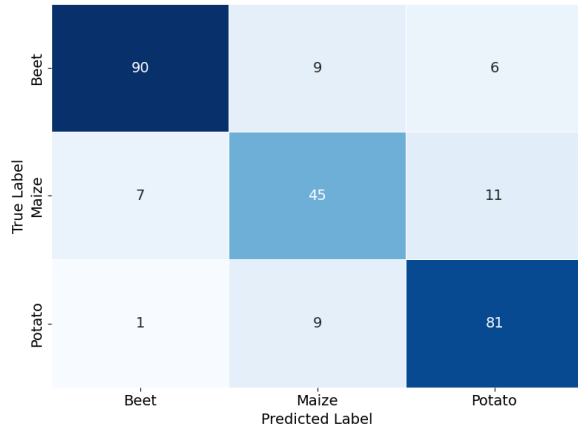
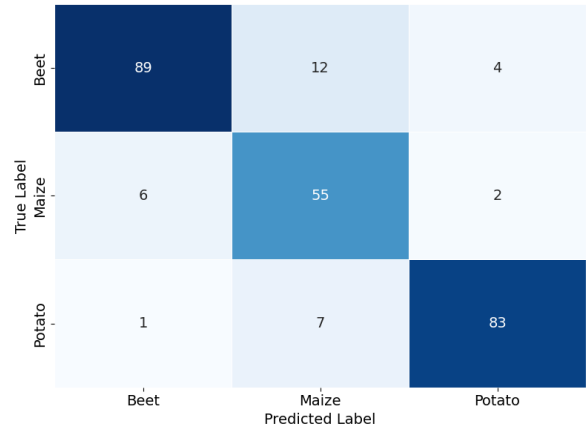


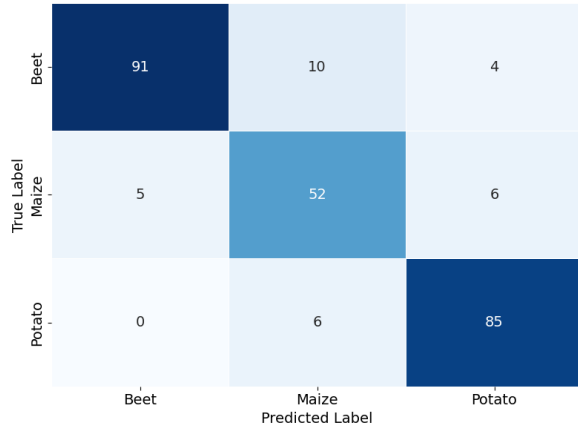
Figure 8.23: Top ten of most important features for the Random Forest classification of beet, maize, and potato crops using Sentinel-1 data on the area of interest, while maintaining a minimum threshold of 150 pixels of image per field.



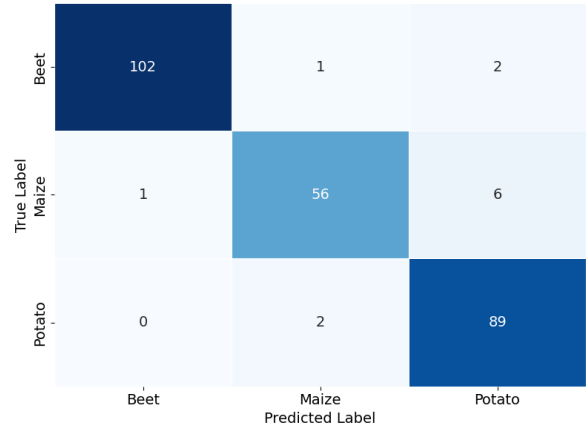
(a) 9th of May



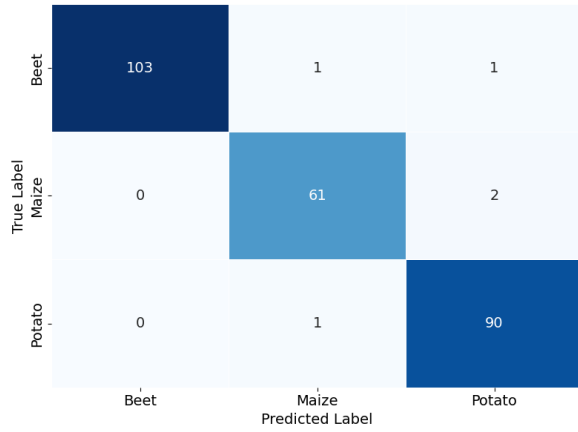
(b) 20th of May



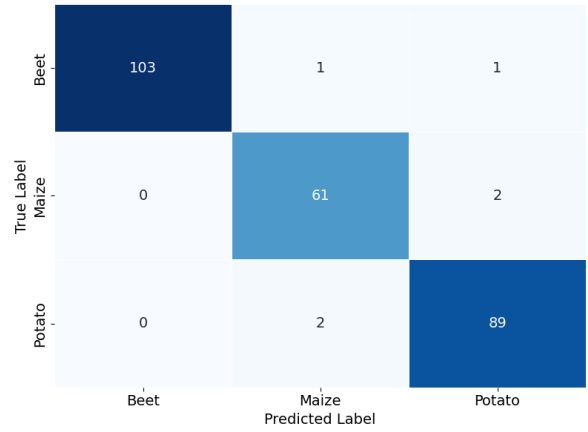
(c) 31st of May



(d) 11th of June



(e) 22nd of June



(f) third of July

Figure 8.24: Last six confusion matrices derived from the Random Forest classification of beet, maize, and potato crops on the area of interest with incremental addition of a combination of TerraSAR-X and Sentinel-1 acquisitions data, while maintaining a minimum threshold of 50 pixels of image per field.

Multi-frequency SAR time series for the detection of seedbed preparation and early-season crop classification

Boland Basile

In the context of the sustainable development goals driving 21st-century debates, agriculture faces the challenge of reducing environmental impact while meeting global consumption needs. This study focuses on the use of multi-frequency SAR imagery, more specifically X-band and C-band data, to evaluate their individual and combined roles in achieving seedbed preparation detection and early season classification for common spring crops cultivated in the southern region of Belgium, including beet, potato, and maize. A seedbed preparation detection method is developed based on the date associated to the minimum VV coherence point between mid-March and late May. These first results are promising and allow to highlight homo- and heterogeneity in spatial and temporal distribution of sowing events between the different cultures. Concerning crop classification, combining C-band and X-band data yields the best results, achieving reliable classification of beet fields in the beginning of April, potato fields by late April, and maize fields by late May. Balanced class distribution, sowing date, and temporal resolution are identified as crucial factors for early classification performance. Seedbed preparation detection and early-season crop classification can serve as valuable tools to build a more sustainable agricultural sector by predicting yield shortenings through agricultural models, providing farmers with helpful management practices and monitor soil preparation activities.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

Faculté des bioingénieurs

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