

Early warning system against muddy floods in central Belgium

Presented by Alice Danzin

Supervisors : Prof. Charles Bielders (UCL/ELI/ELIE)
Prof. Pierre Defourny (UCL/ELI/ELIE)

Readers : Prof. Julien Radoux (UCL/ELI/ELIE)
Mr. Alexandre Mauganrd (UCL/ELI/ELIE)

Mémoire de fin d'études présenté en vue de
l'obtention du diplôme de Bioingénieur :
**Master en sciences et technologies de
l'environnement**

Acknowledgements

The achievement of this master thesis has been a journey and the greatest lesson that I will carry home is that a failure can always be learned from. In this context, I would like to genuinely thank everyone who has participated in the completion of my master thesis. First, I would like to sincerely thank my supervisors, Pr. Charles Bielders and Pr. Pierre Defourny, with whom I have worked with along this year to complete this master thesis and from whom I have learned more than one skill that I am sure will be useful in my future carrier.

Thanks to my readers, Alexandre Maignard and Julien Radoux, for taking the time to read my master thesis and to consider its subject worthy of their attention.

Then, I want to thank Jolan Wolter, researcher in the geomatics laboratory of the Agro faculty, who has been of an invaluable support by answering all of my questions and helping me carry out my experiments. I also want to thank Nicolas Bellemans, researcher and PhD student in the geomatics laboratory of the Agro faculty, who has provided all the data necessary to achieve my goals.

Thanks to my parents, Geneviève Dierge and Didier Danzin who have stood by me through this project and who have participated in the reading of my master thesis. Then, I would like to particularly thank H  l  ne Sosnowski, aka "ma goug", who has been with me through it all, supporting each other and making sure of each other well-being during this experience. I would also like to thank my roommate Aurore de Mahieu for her undying moral support and for taking care of me during the last part of my master thesis. Thanks to my amazing "Eau et sol" team - Jonathan Belis, Tom Busselen, Nathan Bertouille, Aurore de Mahieu, Lucas Demierbe, Charline Onderet, H  l  ne Sosnowski - for standing by each other no matter the obstacle and for the support through theses amazing college years.

Finally, I would like to thank myself for not burning the entire Cérès computer room during the long hours spent trying to carry out my experiments.

Contents

List of Abbreviations	9
Introduction	11
1 State of art	15
1.1 Soil erosion and muddy floods in central Belgium	15
1.1.1 Causes and consequences of muddy floods in Belgium	16
1.1.2 Measures and strategies to reduce muddy floods	17
1.2 Cropping systems in central Belgium (crops, rotations, practices)	20
1.3 Crop and intercrop identification	24
1.3.1 Missions	24
1.3.2 Vegetation indexes	26
1.3.2.1 Normalized Difference Vegetation Index	27
1.3.2.2 Leaf Area Index	27
1.3.2.3 Soil-Adjusted Vegetation Index	27
1.3.2.4 Normalized Difference Water Index	28
1.3.2.5 Brightness index	28
1.3.3 Crop identification methods	28
1.3.3.1 Image processing	31
1.3.4 Examples	33
1.3.4.1 Crop identification using single-date multispectral imagery	33
1.3.4.2 Parcel-based crop identification	35
1.3.4.3 Object-based crop identification	36
2 Methodology	39
2.1 Study zone	39
2.2 Data	40
2.2.1 Land Parcel Identification System map	40
2.2.2 Sentinel-2 products	41
2.3 Softwares	41

2.4	Overall structure of the methodology	42
2.5	Image processing	43
2.6	Crop rotations analysis based on the NDVI time series	44
2.7	LPIS map processing	45
2.8	Rule-based Classification	46
2.9	Maximum Likelihood Classification	47
2.9.1	Pre-classification steps	47
2.9.2	Maximum Likelihood classification methodology	49
2.9.3	Accuracy assessment	50
3	Results and discussion	51
3.1	Crop rotation analysis based on the NDVI time series	51
3.1.1	Basic statistics in 2017	51
3.1.2	Sugarbeet	54
3.1.3	Maize	56
3.1.4	Potato	57
3.1.5	Winter wheat	58
3.1.6	Winter barley	60
3.2	Rule-based Classification	61
3.3	Maximum Likelihood Classification	68
3.4	General discussion	74
	Conclusion	77
	Bibliography	78
	Appendices	85
A	Decision tree used by Peña-Barragán <i>et al.</i> (2011) for the OCIM methodology	87
B	NDVI standard profiles	89
C	Rule-based classification: Confusion Matrices	91
D	Maximum-Likelihood classification: Confusion Matrices	95

List of Figures

1	General patterns of the prediction schemes and its challenges	12
1.1	Frequency of muddy floods over a 10-year period in the municipalities of central Belgium; data for Wallonia: 1991-2000; data for Flanders: 1995-2004 (Evrard <i>et al.</i> , 2007 [16])	16
1.2	Cultural calendar for the main crops identified in central Belgium (Bielders <i>et al.</i> , 2011b)	20
1.3	Calendar of the main winter cover crops in central Belgium (Bertin, 2017)	22
1.4	Spectral signature features of green vegetation	26
2.1	Location of the study zone over the different municipalities	39
2.2	Land use of the study zone, delineated by the blue square (Data source: SPW)	40
2.3	Diagram representing the different steps of the transformation of the LPIS maps into the classification maps	42
2.4	Graphical NDVI evolution of winter wheat, winter barley and winter cover crops (regrouped in one class) between September 2016 and March 2017. The color bands represent the 95% confidence intervals of the mean values	47
2.5	Winter wheat at the end of the tillering stage (Bertin, 2017)	49
3.1	Graphical evolution of the NDVI average of sugarbeet crops in 2017 preceded by winter wheat crops (red), winter barley crops (blue), spring crops (green) or other crops (black) in 2016, showing the 97,5% and 2,5% percentiles around the mean as the coloured patches for each curve	55
3.2	Graphical evolution of the NDVI average of maize crops in 2017 preceded by winter wheat crops (red), winter barley crops (blue), spring crops (green) or other crops (black) in 2016, showing the 97,5% and 2,5% percentiles around the mean as the coloured patches for each curve	57

3.3	Graphical evolution of the NDVI average of potato crops in 2017 preceded by winter wheat crops (red), winter barley crops (blue), spring crops (green) or other crops (black) in 2016, showing the 97,5% and 2,5% percentiles around the mean as the coloured patches for each curve	58
3.4	Graphical evolution of the NDVI average of winter wheat crops in 2017 preceded by spring crops (green), winter cereals (magenta) or other crops (black) in 2016, showing the 97,5% and 2,5% percentiles around the mean as the coloured patches for each curve	59
3.5	Graphical evolution of the NDVI average of winter barley crops in 2017 preceded by winter wheat crops (red), winter barley crops (blue), or other crops (black) in 2016, showing the 97,5% and 2,5% percentiles around the mean as the coloured patches for each curve	60
3.6	Evolution of the mean NDVI averages of winter wheat, winter barley and winter cover crops with their respective standard deviation from August 2016 to March 2017	61
3.7	Graphs representing the kappa evolution of the obtained results in the last rule-based classification tests	66
3.8	Evolution of the NDVI averages of winter cover crops and other crops each followed by maize, sugarbeet or potato, with their respective standard deviation from August 2016 to March 2017	67
3.9	Results of the accuracy assessment of the Maximum Likelihood classification using single-date imagery and two features: NDVI images and 10 bands images	71
A.1	Decision tree for the OCIM methodology developed by Peña-Barragán <i>et al.</i> (2011 [33])	88
B.1	NDVI standard profiles of the spring crops, the winter cereals and the cover crops	89
C.1	Confusion matrices rule-based classification test 1	91
C.2	Confusion matrices rule-based classification test 2	92
C.3	Confusion matrices rule-based classification 3 classes and a parcel-based approach	92
C.4	Confusion matrices rule-based classification 3 classes extended and a parcel-based approach	93
D.1	Confusion matrices ML classification 3 classes (one feature, one date): NDVI images	95

D.2	Confusion matrices ML classification 3 classes (one feature, one date): 10 bands images	96
D.3	Confusion matrices ML classification 3 classes (two features, one date): NDVI & 10 bands images	97
D.4	Confusion matrices ML classification 3 classes (two features, several dates): NDVI & 10 bands images	98
D.5	Confusion matrices ML classification 3 classes (two features, several dates): NDVI, 10 bands & SAVI images	98
D.6	Confusion matrices ML classification 3 classes (two features, several dates): NDVI, 10 bands & NDWI images	99
D.7	Confusion matrices ML classification 3 classes (two features, several dates): NDVI, 10 bands & Brightness images	99
D.8	Confusion matrices ML classification 3 classes (two features, several dates): NDVI, 10 bands, SAVI & NDWI images	100
D.9	Confusion matrices ML classification 3 classes (two features, several dates): NDVI, 10 bands, SAVI & NDWI images	100
D.10	Confusion matrices ML classification 4 classes (two features, several dates): NDVI & 10 bands images	101
D.11	Confusion matrices ML classification 4 classes (two features, several dates): NDVI & 10 bands images	101

List of Tables

1.1	Example of confusion matrix where n_{ij} is the number of samples corresponding to the land cover class map i and validation information j (Defourny, 2017a)	30
1.2	Strengths and weaknesses of algorithms used for classification of satellite imagery (Defourny, 2017a;Autre Gomez)	33
2.1	Main crop rotations identified of the different agricultural regions in Belgium by the GISER Convention (Biielders et al., 2011 [4])	44
2.2	Data assigned to the new fields "Value" and "CLASS"	45
3.1	Basic statistics calculated for 2017 including the type of crops, the number of parcels the percentage of the type of crop compared to the total amount of parcels and the area in hectares	52
3.2	Basic statistics calculated for 2017 including the type of spring crops, the number of parcels, the percentage of the type of crop compared to the total amount of spring crops parcels ($\%_{sp}$), the percentage of the type of crop compared to the total amount of parcels ($\%_{tot}$) and the area in hectares .	52
3.3	Basic statistics calculated for 2017 including the type of winter cereals, the number of parcels, the percentage of the type of crop compared to the total amount of winter cereals parcels ($\%_{wc}$), the percentage of the type of crop compared to the total amount of parcels ($\%_{tot}$) and the area in hectares	52
3.4	Basic statistics calculated for 2017 including the three main spring crops in 2017, what preceded each of them in 2016, the number of parcels and the proportion of each antecedent compared to the total number of parcels of the corresponding spring crop (i.e. maize, sugarbeet, potato)	53
3.5	Basic statistics calculated for 2017 for winter wheat and winter barley in 2017, what preceded each of them in 2016, the number of parcels, the proportion of each antecedent compared to the total number of parcels of the corresponding winter cereal	54
3.6	NDVI thresholds of the four classes for the first test of the rule-based classification method	63

3.7	NDVI thresholds of the three classes used for the second test of the rule-based classification method	63
3.8	Results from the first test of the rule-based classification method using thresholds defined in Table 3.6	64
3.9	Results from the second test of the rule-based classification method using thresholds defined in Table 3.7	64
3.10	NDVI thresholds of the three classes for the additionnal dates	65
3.11	Results of the parcel-based test for all the dates previously identified for the three classes and the three classes using the extended parcel data set	65
3.12	Types of tests in the Maximum Likelihood Classification	68
3.13	Results from the Maximum Likelihood classification method (0); UA and PA are of the WCC class (25/09/2016, 04/12/2016, 27/12/2016, 15/02/2017, 14/03/2017)	69
3.14	Results from the Maximum Likelihood classification method (1); UA and PA are of the WCC sub-classes (04/12/16 & 27/12/16 & 15/02/17 & 14/03/17)	72
3.15	Results from the Maximum Likelihood classification method (2); UA and PA are of the WCC class	73
3.16	Comparison of the kappa obtained using three classes and a pixel-based approach in rule-based and maximum likelihood classification	74

List of Abbreviations

AEM = Agro-Environmental Measures

CAP = Common Agricultural Policy

DT = Decision Tree

ESA = European Space Agency

GISER = Gestion Intégrée Sol Erosion Ruissellement

HRVIR = High Resolution Visible and InfraRed

LAI = Leaf Area Index

LPIS = Land Parcel Identification System

MIR = Mid-InfraRed

ML = Maximum Likelihood

MSS = Mutlispectral Scanner System

NASA = National Aeronautics and Space Administration

NDVI = Normalized Difference Vegetation Index

NDWI = Normalized Difference Water Index

NIR = Near InfraRed

OA = Overall Accuracy

OBIA = Object-based Image Analysis

PA = Producer Accuracy

PGDA = Programme de Gestion Durable de l'Azote

RF = Random Forest

SAVI = Soil-Adjusted Vegetation Index

SPW = Service Public de Wallonie

SWIR = Short Wave InfraRed

UA = User Accuracy

USGS = United States Geological Survey

Introduction

Muddy floods are recurrent and increasing problems in many regions throughout Europe (Boardman *et al.*, 2010). They arise when intense rains fall on bare or poorly covered land, generating intense runoff and soil erosion. Depending on the landscape patterns, water loaded with sediments may be channeled towards other fields or towards public and private infrastructures causing extensive damage (Evrard *et al.*, 2007b).

In central Belgium, muddy floods mostly occur in spring and early summer (Evrard *et al.*, 2007b). This results from a combination of highly erosive rainfall events and the circumstance of large contiguous surfaces of spring crops (e.g., sugarbeet, maize, potato) which provide poor soil cover at that time of the year (Evrard *et al.*, 2007b). Additionally, in central Belgium, as the surface area of spring crops increases every year, the risk of muddy floods rises as well. Muddy floods are further facilitated by the silt loam soils which are prone to surface crusting, thereby enhancing runoff and erosion.

The cost of these muddy floods has been estimated by Evrard *et al.* (2007a) at 16 to 172 million euros per year in central Belgium. Because of the high costs involved and the psychological stress that affects people subjected to muddy floods, measures need to be taken in order to reduce their occurrence and the magnitude of their consequences. Although soil conservation measures aiming at reducing the occurrence of runoff and erosion would be preferable, implementing such conservation measures over extensive areas will require long lasting efforts. In addition, they may not be fully effective in case of extreme rainfall events. Alternatively, measures can be taken to minimize the downstream impact of muddy flows by retaining sediment upstream as much as possible. Grass strips, but also porous filters such as fascines and straw filters can be used to that purpose. However, the number of potential muddy flood sites is very high. And, the location of muddy flood-affected sites will differ from year to year since the spatial distribution of fields cropped with spring crops is highly variable. Hence, it would be advantageous to be able to predict three or six months in advance the areas most at risk of muddy floods.

Central to the success of a muddy flood early warning system would be the ability to pre-

dict which cropland fields will be bare or poorly covered in spring and early summer, i.e., which fields will be cropped with spring crops in year $n+1$. In central Belgium, the main crops are sugarbeet, potato and maize (spring crops) whose growing cycles fall between early spring and late autumn, and winter barley and winter wheat (winter cereals) whose cycle fall between autumn and the following summer. One way to predict the type of crop that will be sown in spring of year $n+1$ is to rely on knowledge regarding the most common crop rotations as well as knowledge of the previous crop(s) in year(s) n , $n-1$, etc (Figure 1). Typical crop rotations can be inferred from the Land Parcel Identification System (LPIS) which combines the land area declarations that Belgian farmers have to fill every year to specify which crop they plant on each plot of land. However, with time, the diversity of crop rotations has increased, making it increasingly difficult to predict a future crop on the basis of the previous crop(s) (Urdiain-Arraiza, 2018). The LPIS also allows determining the crop in years n , $n-1$, ... Even though the information to build the LPIS is gathered each year, crop type of year n is made public only at the beginning of year $n+1$, which constitutes a further restriction for its use. Alternatively, one could use the Sen2Agri crop type product for a timely determination of the type of crop in year n , but this information would necessarily be less precise than the LPIS and would not solve the problem of low predictability of crop rotations.

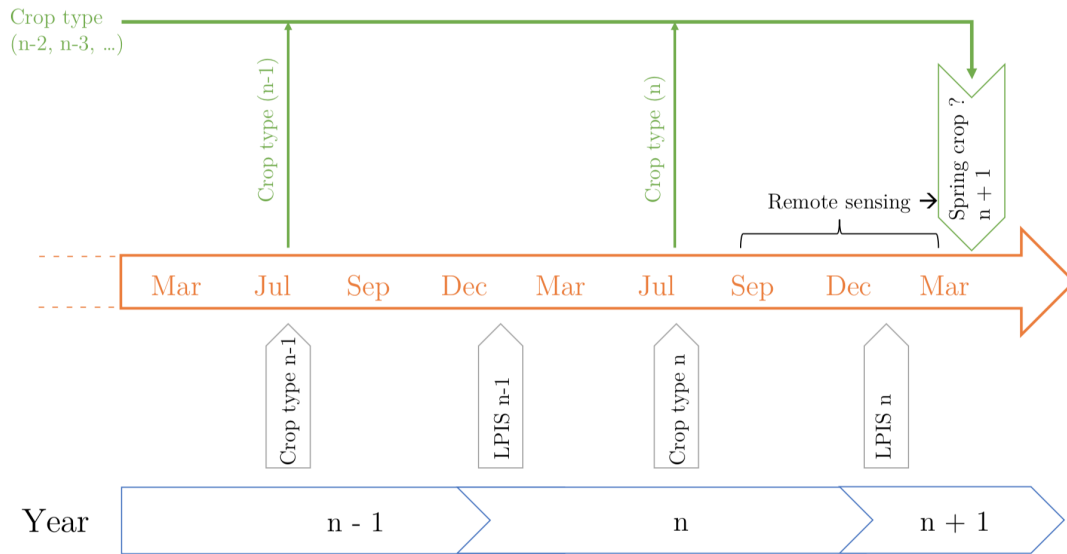


Figure 1: General patterns of the prediction schemes and its challenges

An alternative to the use of LPIS and crop rotations resides in the determination of soil cover during late autumn – winter period. Indeed, plots not sown with winter cereals would have a high probability of being sown with spring crops in the spring. Remote

sensing offers a unique opportunity to monitor land cover. With the new Sentinel-2A and Sentinel-2B satellites, land cover may be derived at a high spatial resolution and with high temporal frequency, thereby facilitating the identification of the cover type at plot scale. However, during the months preceding planting, the plots destined for spring crops may be covered with late-harvested spring crops (e.g., sugarbeet), various winter cover crops with highly variable development stages, or may be left bare (tilled or not). This makes for highly heterogeneous conditions. Therefore, the challenge of predicting the crop type in spring will be to distinguish between these highly variable surface conditions on the one hand, and winter cereals on the other hand, during the preceding autumn-winter period (see Figure 1).

Finally, the aim of this master thesis is to evaluate the accuracy with which one can predict the occurrence of spring cropped fields (sugarbeet, potato, maize) based on land cover information acquired prior to sowing of spring crops. To achieve this purpose, the first task will be to analyze vegetation indexes time series acquired from data retrieved by satellites. Then, temporal windows will be identified in order to determine key dates to distinguish winter cereals from winter cover crops. Then, satellites images will be classified and an assessment of the classifications will be performed in order to locate areas at risk of muddy floods during spring and summer.

The first chapter of this master thesis will consist in a the state of art reviewing the general patterns of muddy floods in central Belgium, the calendars and growth cycles of the main crops cultivated in this region and an overview at the techniques already developed to identify and classify crops using remote sensing. Then, in the materials and methods chapter, the methodology will be presented. After exposing the different classification methods tested in this thesis the results will be discussed before concluding the overall achievements of this thesis.

Chapter 1

State of art

1.1 Soil erosion and muddy floods in central Belgium

Soil erosion is the process by which the soil and relief are altered through detachment and transport of soil particles by various means : water, wind, cultural practices (tillage, harvest). There are different types of erosion : water erosion, wind erosion, harvest erosion, tillage erosion, mass transport.

Water erosion processes are the most important in Belgium, especially in the central part of the country. Water erosion processes in Belgium are mostly significant on agricultural plots. Water erosion has many consequences. First of all, the destruction of the soil structure. As the rain falls, raindrops crush the soil aggregates and create a hard soil crust (or slaking crust) that restricts water infiltration through the soil surface. Then, water erosion generates many losses: soil, nutrient, organic matter and fertilizer losses. Furthermore, water erosion may also cause damage to the crops by exposing freshly sown seeds and plant roots in eroding soils and by covering seedlings in sediment deposition areas. This mostly results in economical damage. Furthermore, runoff loaded with sediment from agricultural parcels may cause muddy floods, as observed in many Belgian municipalities and especially in the Belgian loam belt (Evrard *et al.*, 2007a).

Muddy floods are "off-site" consequences of water erosion. They mainly occur after intense rain events especially in spring and in the beginning of summer and generally originate from cultivated areas. Muddy floods depend on the meteorological event and the landscape vulnerability which is deeply connected to cropping practices and cropland surface (Evrard *et al.*, 2007a).

1.1.1 Causes and consequences of muddy floods in Belgium

The location of municipalities influences the magnitude of muddy floods. As a matter of fact, municipalities situated upstream of main rivers are more prone to endure direct runoff from agricultural parcels in the form of muddy floods while cities located near large rivers, more specifically in alluvial plains, suffer more from localized flooding of these streams (Verstraeten *et al.*, 1999).

As stated by Bielders *et al.* (2003), 27% of wallonian municipalities have suffered from muddy floods at least once every two years between 1990 and 2000. In addition, Verstraeten *et al.* (1999) set up a questionnaire reaching out to 137 municipalities in central Belgium to assess the magnitude of the muddy floods issue: 43% have to deal with muddy floods at least once a year. For this reason, Evrard *et al.* (2007a) carried out a survey to quantify off-site consequences of muddy floods. Hence, Figure 1.1 shows the frequency of muddy floods in central Belgium over a 10-year period. This research concludes that 79% of the 201 municipalities studied face at least one muddy flood every ten years. Besides, 22% of these impacted municipalities have experienced more than ten muddy floods over the 10-year period.

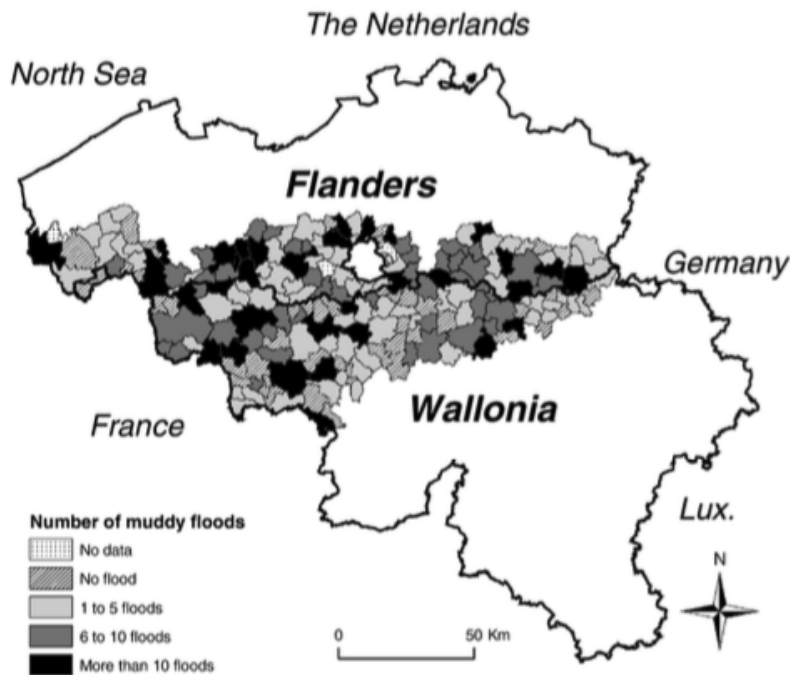


Figure 1.1: Frequency of muddy floods over a 10-year period in the municipalities of central Belgium; data for Wallonia: 1991-2000; data for Flanders: 1995-2004 (Evrard *et al.*, 2007 [16])

Furthermore, in the last few years, a lot of municipalities in central Belgium have been

seriously damaged by muddy floods both economically and socially. According to Evrard et al. (2007a), the frequency of muddy floods has increased significantly in central Belgium, due to the expansion of spring crops at the expense of winter cereals and grassland (Evrard et al., 2007b). This statement is also corroborated by Biëlders et al. (2003) who stipulate that European policies favored the increase of crops such as sugarbeet, potato, chicory, which are transformed industrially. These spring crops (or summer crops) require wide spacing and a thin seedbed for the plants, thereby unfortunately enhancing surface sealing and reducing its roughness. Thus, these processes raise runoff volume and velocity leading to the increase of peak discharge (Evrard et al., 2007a). Moreover, spring crops cover the soil surface insufficiently during May and June reducing the soil protection when rainfall erosivity is highest. Rainfall erosivity expresses the ability of the rain to induce runoff depending on the intensity and the duration of the rain and its capacity to detach the soil particles (kinetic energy). Hence, rainfalls in May and June are prone to generate muddy floods quickly even though they may not be extreme. In addition, the field size enlargement in order to increase productivity and the construction of new roads to access the fields may also explain the intensification of muddy floods (Evrard et al., 2007b).

The risk of muddy floods is a combination between the vulnerability of private and public infrastructures and the capacity of a meteorological event to generate runoff. The landscape sensitivity also represents an important aspect of that risk (Verstraeten *et al.*, 1999). Mitigation measures can be taken regarding landscape vulnerability. However, the basic driving force of muddy floods, which is the rain, cannot be altered by human actions (Verstraeten *et al.*, 1999).

1.1.2 Measures and strategies to reduce muddy floods

The cost of muddy floods has been assessed in a survey carried out by Evrard et al. (2007a). The cost for cleaning and repairing damage to public infrastructures has been evaluated to 12,5 to 122 million euros per year in central Belgium. Similarly, the cost of damage to private property has been estimated from 1,6 to 16,5 million euros per year. Given the yearly huge costs resulting from muddy floods and the psychological stress that affects people due to these events, actions must be taken. Therefore, as a result of the necessity to reduce muddy flood damage and their costs, farmers, municipalities and local people may come up with different strategies to achieve this purpose.

First of all, mitigation measures originating from the European Common Agricultural Policy¹ can be implemented through the Walloon public service (SPW). In order to be

¹CAP: agricultural policy created by the European Commission

eligible for their subsidies, farmers are bound to some standards. In Wallonia, fields with a slope steeper than 10 per cent in half of their surface are submitted to specific rules. Alternative agricultural practices must be applied on those fields. Concerning Agro-environmental Measures (AEM's), farmers have the opportunity to set up these measures in order to benefit from European subsidies and the environmental services they offer (Evrard *et al.*, 2010).

Above all, two measures contribute to reduce runoff and diminish erosion: grass band strips and cover crops (Evrard *et al.*, 2010). The first one is an AEM and the second one is an obligation for the farmers as part of the Nitrate Directive. Grass buffer strips consist in grassy areas planted at the edge of cropped fields. They are composed of grasses and legumes. Their effect on muddy flood mitigation is the protection of the soil and the sedimentation of soil particles contained in surface runoff (Natagriwal). The farmer receives 21,6 euros per twenty meters long strip and at least twelve meters wide. Then, cover crops are planted between a winter cereal and a spring crop, during the so-called: "dormant period". The details about cover crops and the Nitrate Directive will be elaborated in Section 1.2.

In order to help farmers and municipalities manage muddy floods risks, the Wallonia Public Service (SPW) set up a research and advice unit called GISER² (Integrated Management Soil Erosion Runoff) that provides technical council and urban planning advice for erosion management. With the aid of this unit, municipalities can set up strategies to deal with runoff, soil erosion and muddy floods in cooperation with farmers (Janssens *et al.*, 2012).

Furthermore, municipalities may implement local scale measures. First, retention ponds are the most widespread retention structures in central Belgium, and even in the entire country (Evrard *et al.*, 2010). Retention ponds are designed to collect water-runoff from violent rainfall events which may transport a large amount of sediments. They also reduce peak flows. However, as sediments gradually fill the retention pond, its capacity is diminished leading to the disability to store water during successive extreme events (Verstraeten *et al.*, 1999). Therefore, the maintenance of these structures is essential. Verstraeten *et al.* (1999) evaluated the construction cost of retention ponds at 380 000 euros, which is very expensive. They also estimated the cost of removing one cubic meter of sediments at 8 to 10 euros, not to mention the costs of the structures maintenance. So, the implementation and maintenance costs of mitigation measures might be high but the overall cost of muddy floods is higher. Therefore, coming up with strategies to reduce

²Gestion Intégrée Sol Erosion Ruissellement

the impacts of muddy floods is essential to prevent damage, economical loss and psychological stress.

Finally, since the magnitude of muddy flood consequences is tremendous, setting up a muddy flood early warning system represents a cost-effective approach to reduce the risks of damage. According to the European Climate Adaptation Platform, to be efficient an early warning system should consist in four elements: risk knowledge, monitoring and warning service, dissemination and communication, and response capability (European Climate Adaptation Platform). In this thesis, given that the main risk factors are well known in Wallonia, the main focus is on the monitoring part.

1.2 Cropping systems in central Belgium (crops, rotations, practices)

In Belgium, there are ten different agricultural regions and each region has its specialties. For example, the loam region is known as a field crop region and the Ardennes is known as a livestock breeding region. Central Belgium is highly dominated by the loam region. This region, known as a field crop region, has five major crops : winter wheat, sugarbeet, potato, winter barley and maize (Bielders *et al.*, 2011). Leteinturier *et al.* (2007) carried out a study to make the agronomic and environmental diagnosis of crop rotations in Wallonia. According to this research, between 1997 and 2003, winter wheat was the most represented major crop in the loam region, covering 38% of the total surface of major crops which accounted for 142 166 hectares. In 43% of this cases, winter wheat was preceded by sugarbeet. During this period, the second most important major crop was sugarbeet covering 21% of the total surface previously mentioned. This kind of crop was preceded by winter wheat in 53% of the cases analyzed. And finally, the third most important major crop was silage maize³ covering the total surface at 9% and was preceded by winter wheat in 43% of the cases.

Figure 1.2 shows a crop calendar based on the information gathered by Bielders *et al.* (2011b).

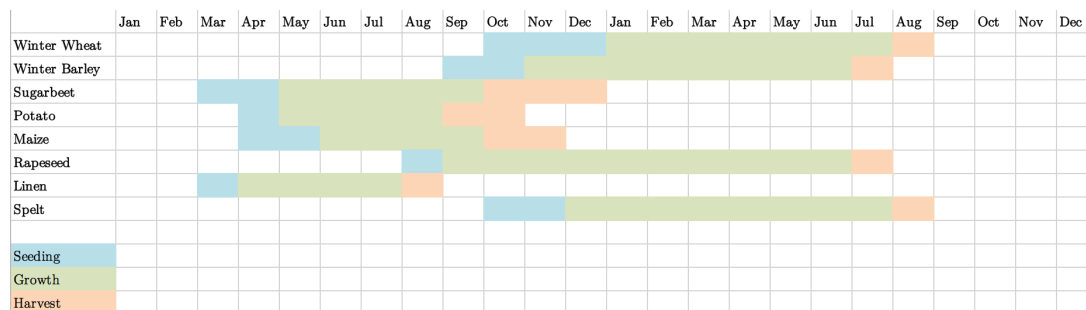


Figure 1.2: Cultural calendar for the main crops identified in central Belgium (Bielders *et al.*, 2011b)

In central Belgium, a spring crop is often preceded by a winter cereal. As winter cereals are harvested in July and August, there is a gap until the spring crops are sown in March-May of the next year. The common trend consists in planting winter cover crops (also known as intermediate crops) which will cover the soil in winter and provide various benefits for the spring crop: providing nutrients for the spring crop, preventing runoff, improving soil structure, etc. Thus, identifying winter cover crops represents an

³maïs ensilage

important aspect of the prediction scheme in order to acquire the ability to distinguish these crops from winter cereals.

For each major crop, the management of the intercrop period is different depending on the next crop. After a winter cereal, like winter wheat or winter barley, there are two options. On the one hand, the following crop is another winter cereal. The crop residue is left on the plot of land and the "new" cereal is sown shortly after the first cereal's harvest. On the other hand, the winter cereal can be followed by a spring crop (like sugarbeet, potato, maize, etc). In that case, the farmer can choose between three options. First, an intermediate crop can be planted after the winter cereal's harvest, generally in August-September. The intermediate crop can either:

- stay on the plot until frost destroys it, or can be destroyed using glyphosate or by mechanical processes which depends on the species;
- be destroyed by the farmer and buried from November 15.

Then, the spring crop is sown in March or later on. Second, the crop residue is left on the parcel until the land is plowed in autumn (November). Finally, the last option is to leave the crop residue on the plot and till the soil in spring (March).

In Belgium, sugarbeet is a very important crop. The harvest may begin in October and can last until late December or January depending on the percentage of sugar in the beet. A sugarbeet crop can be followed by a winter cereal: winter wheat or spelt, if it is harvested in November. If not, it will be too late for the winter cereal to grow. Alternatively, after a sugarbeet crop, another spring crop can also be planted. Either the plot is plowed right after the sugarbeet's harvest and left bare until the following spring crop is sown. Or, the sugarbeet residue are left on the parcel after the harvest, the soil is tilled in the spring and after that the next spring crop can be planted.

Many agricultural regions produce potatoes because it is a very popular foodstuff in Belgium. Generally, potatoes are harvested at the end of September and until mid-October. Shortly after that, a winter cereal can be sown. However, the farmer can choose to plant another spring crop after the potatoes have been harvested. Then, the potato residue is left on the parcel and there can be an autumn (November) or a spring (March) tillage. The soil is left bare after the tillage until the following spring crop is sown.

There are two types of maize cultivated in Belgium: silage maize intended for bovine consumption, and grain maize which serves as feed rations for pigs, poultry and concentrates for bovines. Maize is usually harvested in mid-October. After that, the farmer can

sow a winter cereal like winter wheat or spelt. But there is another option: the maize crop will be followed by another spring crop. So, the maize residue is left on the parcel until the land is plowed in autumn (November) or in spring (March). Then, the soil is left bare before the next spring crop is planted (Bielders *et al.*, 2011a).

As previously mentioned, winter cover crops can be planted between two successive main crops of the rotation. This is also called: intercropping. The conventional example of intercropping is between winter barley and spring crops (sugarbeet, maize, potato, ...). But this practice is also applied after peas and rapeseed. Intercropping pursues several purposes:

- It may be used as green manure once the fields are tilled;
- It also serves as soil cover in order to fight against soil erosion;
- It serves as a nitrate catch crop.

A lot of plant species can be used as winter cover crops depending on their purpose and the farmer's intention. Figure 1.3 gathers some of the most used winter cover crops in central Belgium and also shows their growth calendar.

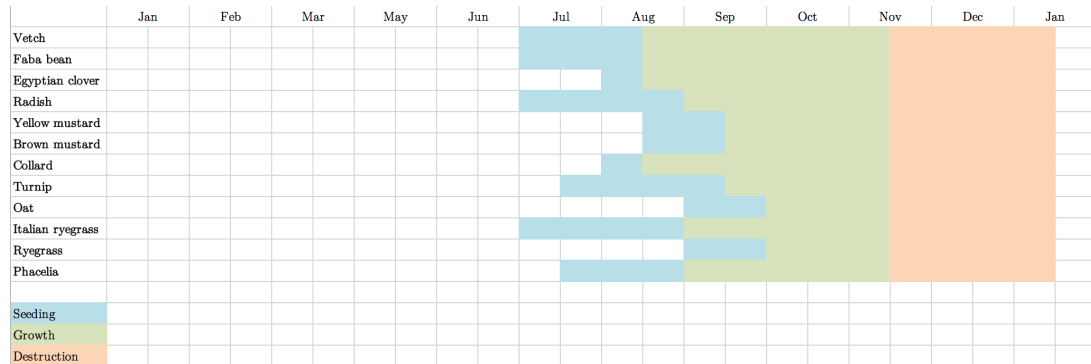


Figure 1.3: Calendar of the main winter cover crops in central Belgium (Bertin, 2017)

In Belgium, some rules regulate the use of winter cover crops. They depend on the location of agricultural plots. These rules are applied according to the Nitrate Directive⁴ which is part of the Sustainable Management Programme of Nitrogen (PGDA⁵) in agriculture, implemented since June 2014. As a matter of fact, Wallonia is divided into two zones agriculturally speaking: a vulnerable zone and a non-vulnerable zone. First, in the non-vulnerable zone, if there is an input of organic amendements between the 1st of July and the 15th of September, plant covers have to be sown before the 15th of September

⁴Directive Nitrate

⁵Programme de Gestion Durable de l'Azote

included and they can be destroyed from the 15th of November. From that moment, farmers can either destroy the winter cover crop or let the frost destroy it. Then, in the vulnerable zone, 90% of the harvested zones before the 1st of September and that will be a spring crop the next year must be covered by a winter cover crop by September 15, which has to be maintained until the 15th of November. Winter cover crops can be destroyed from the 15th of November. In the vulnerable zone, for every legume crop harvested before the 1st of August and followed by winter wheat, the rule states that winter cover crops have to be sown before the 1st of September and destroyed from the 1st of October (Nitrawal; PROTECT'eau).

As winter cover crops grow when winter cereals also start to develop, the challenge of this master thesis will be to find the patterns that allow the distinction between these two types of crops according to the growth calendars, vegetation features, etc.

1.3 Crop and intercrop identification

As mentioned earlier, the main objective of this master thesis consists in distinguishing winter cereals from winter cover crops. To achieve this purpose, data retrieved by remote sensing will be used and analyzed. The development of remote sensing has been possible throughout the years thanks to the programs launched by different space agencies and organizations. In this section, the Landsat and SPOT programs will be presented. Providing continuity to these two programs, the Copernicus program will also be presented. Then, the vegetation indexes used in this master thesis will be introduced. Finally, identification methods already studied will be presented in order to gain a better understanding of crop identification techniques and their application range.

1.3.1 Missions

Landsat Program

The Landsat program was the first aerospace mission, created by the NASA⁶ and the USGS⁷, with the purpose to observe the earth. Landsat 1 (the first satellite) was launched in 1972 with a camera and a Multispectral Scanner System (MSS) of four bands on board. For over 40 years, the Landsat program has placed eight satellites in orbit and has delivered valuable information about the Earth. The last satellite, Landsat 8, was launched in 2013 and still covers the globe: it takes sixteen days to produce the whole coverage of the Earth. Landsat 8 has eleven spectral bands: eight spectral bands at 30 meters spectral resolution, one panchromatic spectral band at 15 meters spatial resolution and two spectral bands in the thermal infrared at 100 meters spatial resolution (Defourny, 2017).

SPOT VEGETATION

SPOT⁸ (Probative Observation System of the Earth) is an aerospace program created by a partnership between three countries: Belgium, France and Switzerland. The first satellite (SPOT 1) was launched in 1986. SPOT 1,2,3 had an HRVIR⁹ instrument on board with three spectral bands at a 20-meter spatial resolution and a panchromatic band at a 10-meter spatial resolution. Then, SPOT 4 and 5 were launched respectively in 1998 and 2002. They both carry measuring instrument HRVIR and MIR¹⁰. Moreover, they also carry the instrument VEGETATION which has a 1-kilometer spatial resolution and a swath

⁶National Aeronautics and Space Administration (USA)

⁷United States Geological Survey

⁸Système Probatoire d'Observation de la Terre

⁹High Resolution Visible and InfraRed

¹⁰Mid-InfraRed

of 2200 kilometers. Thanks to the VEGETATION instrument, SPOT 4 and 5 cover the Earth globally and daily. Another asset of the program is that the panchromatic band spatial resolution has been reduced over the years, going from 10 meters to 5 meters between SPOT 4 and SPOT 5, from 5 to 1.5 meters between SPOT 5 and SPOT 6 & 7 (Defourny, 2017).

The Copernicus program: Sentinel-2

As the main part of the thesis is to distinguish different types of crops from others in a specific zone using remote sensing, the Copernicus program appears to be providing a lot of information regarding the vegetation behaviour. This program was created by the European Space Agency in order to observe the earth and its components. The Copernicus program includes six missions, the Sentinel missions, and each mission is composed of a constellation of two satellites. Moreover, each mission is specialized in a domain among land, ocean and atmospheric monitoring (ESA1).

The Sentinel-2 mission gathers information that is mainly used in land monitoring. It includes twin polar-orbiting satellites that operate simultaneously, located in the same orbit but 180° apart to ensure optimal coverage and data delivery. They are situated in a sun-synchronous orbit at an altitude of 786 kilometers. Their orbital swath width is 290 kilometers. These two satellites monitor variability in land surface conditions.

The mission also provides multispectral high-resolution optical imagery. Its high revisit frequency offers an opportunity in agriculture monitoring. As a matter of fact, the revisit time is 10 days at the equator with one satellite and 5 days with two satellites and cloud-free conditions. So, at mid-latitudes, the revisit frequency is about 2 or 3 days. Therefore, it allows to monitor the changes of vegetation during the growing season and thus identify the crops at risk. In addition to the high revisit time, Sentinel-2 satellites also provide opportunities in plant behaviour monitoring due to their spatial resolution. The satellites carry an optical instrument payload that samples 13 bands: four bands at 10 meters, six bands at 20 meters and three bands at 60 meters spatial resolution (ESA2). The Sentinel-2 mission main objectives are to provide:

- "systematic global acquisitions of high-resolution, multispectral images allied to a high revisit frequency;
- continuity of multi-spectral imagery provided by the SPOT series of satellites and the USGS LANDSAT Thematic Mapper instrument;
- observation data for the next generation of operational products, such as land-cover maps, land-change detection maps and geophysical variables" (ESA2)

Then, two products are provided by Sentinel-2. The Level-1C product: Top-Of-Atmosphere reflectances in cartographic geometry. And the Level-2A product: Bottom-Of-Atmosphere reflectances in cartographic geometry. These products are a compilation of elementary granules, called tiles which size is 100x100 km² ortho-images in UTM/WGS84 projection. The Sentinel-2 mission is furthermore highly compatible with the Landsat missions that offer new opportunities for regional to global agriculture monitoring.

As previously mentioned, Sentinel-2 spatial resolution, revisit frequency, global coverage and its compatibility with the Landsat and SPOT programs provide a wide range of new opportunities from regional to global agriculture monitoring. In this context, the ‘Sentinel-2 for Agriculture’ project has been launched in partnership with the catholic University of Louvain. The project leads to four products: a cloud-free reflectance composite, a cropland mask, a crop type map and vegetation status (Sen2Agri).

In this master thesis, the Sentinel-2 for Agriculture products that will be used are the vegetation status images. Since Sentinel-2A was launched in June 2015 and Sentinel-2B in March 2017, there already is quite an amount of data available. (ESA3).

1.3.2 Vegetation indexes

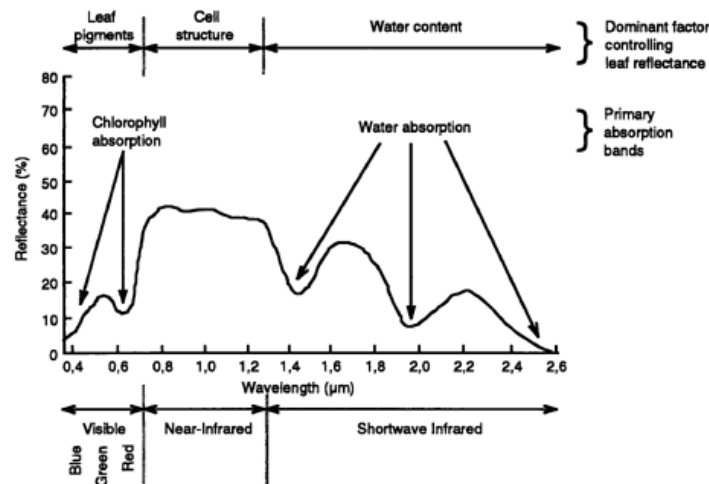


Figure 1.4: Spectral signature features of green vegetation

In every plant, chlorophyll absorb the sun’s energy to perform photosynthesis. In the visible spectrum, leaves absorb red and blue light, and transmit the green light. That’s the reason why leaves are green. But the infrared spectrum also interacts with the vegetation (Figure 1.4). The radiation in the near infrared is able to pass through the plant

cuticle and interact with the spongy parenchyma. Indeed, the infrared radiation is highly reflected inside the leaf and refracted outside the leaf thanks to the intercellular space which is made of air and water. Therefore, observing the plant behaviour in both spectra gives significant clues about the plants characteristics. That is the reason why vegetation indexes are used, because they are based on the difference of reflectance between the visible spectrum and the infrared spectrum (Defourny, 2017).

1.3.2.1 Normalized Difference Vegetation Index

The Normalized Difference Vegetation Index (NDVI) is the most useful index because it characterizes the health state of crops, identifies the phenological changes, allows to assess the green biomass and the yield of the crops (Ledant, 2007). The index was first used by Rouse *et al.*, in 1973.

It's the ratio of the difference between the reflectance in the near infrared (NIR) and the red, and the sum of both. It varies from -1 to +1 :

$$NDVI = \frac{NIR - red}{NIR + red} \quad (1.1)$$

The NDVI rises during the growing season. Thus, a higher NDVI reflects a higher photosynthetic activity. Generally, the NDVI values for a crop range from 0,1 to 0,9 (Quadratic SPRL, 2016).

1.3.2.2 Leaf Area Index

The following useful vegetation index is called Leaf Area Index, LAI, and it analyses the plant canopy. It represents the one-sided green leaf area per unit of ground surface area. It varies from 0, corresponding to bare soils, to 10, corresponding to really dense conifers. In general, LAI values for a crop vary between 0 and 4.

1.3.2.3 Soil-Adjusted Vegetation Index

In 1988, Huete carried out a study in order to present a process in order to minimize the impact of soil brightness on the spectral vegetation indexes involving red and NIR spectral bands. Its purpose is to rectify the influence of the soil brightness when the vegetation cover is low. It is actually a modification of the NDVI:

$$SAVI = \frac{NIR - red}{NIR + red + L} * (1 + L) \quad (1.2)$$

L is the soil brightness corrector factor. When there is no vegetation, L=0 and SAVI = NDVI. In most cases, its value is set at 0,5 (Huete, 1988).

1.3.2.4 Normalized Difference Water Index

The Normalized Difference Water Index is the index derived from remote sensing related to liquid water. It is mostly used to monitor variability in water content of leaves. The following equation was proposed by Gao (1996) who set up a research designing this index to assess the vegetation liquid water content using remote sensing due to its important applications in agriculture and forestry.

$$NDWI = \frac{(NIR - SWIR)}{(NIR + SWIR)} \quad (1.3)$$

1.3.2.5 Brightness index

$$Brightness = \sqrt{Green^2 + Red^2 + NIR^2 + SWIR^2} \quad (1.4)$$

1.3.3 Crop identification methods

With the evolution and development of technologies involved in agricultural management, the need for crop distribution information at plot scale has increased and there have been many studies using a variety of methods to identify different crops (Yang *et al.*, 2017).

Crop identification based on remote sensing is constrained by the selection of the data source which depends on the quality and the spatial detail of the mapping output (Defourny, 2017a). Four main criteria should be taken into consideration regarding land cover mapping. The spatial resolution is the most important criterion. It represents the smallest land feature to be detected and then possibly mapped (Defourny, 2017a).

The frequency of valid cloud-free observations is the second most important criterion. Indeed, the growth cycles of crops and the seasonal dynamics of agricultural parcels are only captured when it is possible to acquire cloud-free time series. So, distinguishing one crop from another does not only requires consideration of the revisit frequency of satellites, it also includes the frequency of valid cloud-free images (Defourny, 2017a). In cloudy regions, constellations of satellites are really useful to reduce the effect of clouds. These constellations also allow distinguishing crops that have very similar growth cycles. As a matter of fact, a lot of crop cycles are comparable because of their growth calendar and their features. But, there is also variability within the same crop type, i.e. crop calendar and features can be different for the same crop type, but also from one plot to another or from year to another (Peña-Barragán *et al.*, 2011).

The third criterion is the number of spectral bands and their position in the electromag-

netic spectrum. This not only important for crop discrimination but also for atmospheric corrections. Nowadays, visible and NIR bands are available on most satellites. The addition of SWIR bands is also very useful in identifying crops (Defourny, 2017a)

Finally, swath width is an important criterion in large-scale applications of the agricultural landscape mapping process. As satellites revolve around the Earth, only a portion of the Earth is "captured", this area is referred as the swath. Depending on the satellite in question, its orbit and the rotation of the Earth, the swath can vary from tens to hundreds of kilometers wide (Defourny, 2017a).

Regarding resolution, one of remote sensing's biggest difficulties is to provide high spatial and high temporal resolution at the same time (Yang *et al.*, 2017). The best compromise is to combine spatial, temporal and spectral resolution: the pixel size must smaller than the agricultural parcels size in order to get a sufficient number of pure cropland pixels (Defourny, 2017a), images at different dates of the crop growth and information at different regions of the optical spectrum (visible, infrared) are needed (Peña-Barragán *et al.*, 2011).

Every land cover mapping process needs to go through calibration followed by validation. The purposes of calibration consists in covering every situation possible existing within the study zone in order to get a sample of the different possibilities of spectral and temporal signatures. The training data set used to calibrate the land cover map will link up the different land cover classes with the existing range of corresponding spectral and temporal signatures. After that, the validation step provides a statistical analysis of the accuracy of the land cover map based on an independent information source.

The accuracy of the map is assessed by evaluating the degree of agreement between the land cover map and the validation data set. The accuracy assessment can be performed by various softwares like Arcmap and QGIS by computing a confusion matrix which provides various metrics of classification accuracy. The accuracy assessment by computing confusion matrices can also be performed for the calibration data set and the validation data set with the purpose to evaluate the quality of both processes. Table 1.1 displays an example of a confusion matrix with the aim of describing some performance metrics. The confusion matrix (Table 1.1) is set to match the number of samples of a given land cover class with the validation data set. Diagonal values represent the agreement between the two data sets whereas the non-diagonal values show the errors.

Table 1.1: Example of confusion matrix where n_{ij} is the number of samples corresponding to the land cover class map i and validation information j (Defourny, 2017a)

	Validation data					
Map	Cropland	Forest	Grassland	Urban	Water	Total
Cropland	n_{11}	n_{12}	n_{13}	n_{14}	n_{15}	$n_{1.}$
Forest	n_{21}	n_{22}	n_{23}	n_{24}	n_{25}	$n_{2.}$
Grassland	n_{31}	n_{32}	n_{33}	n_{34}	n_{35}	$n_{3.}$
Urban	n_{41}	n_{42}	n_{43}	n_{44}	n_{45}	$n_{4.}$
Water	n_{51}	n_{52}	n_{53}	n_{54}	n_{55}	$n_{5.}$
Total	$n_{.1}$	$n_{.2}$	$n_{.3}$	$n_{.4}$	$n_{.5}$	$N_{..}$

In Table 1.1, the row totals correspond to the number of samples classified in category i according to the validation data:

$$n_{i.} = \sum_{j=1}^k n_{ij} \quad (1.5)$$

Then, the column totals refer to number of samples classified into category j according to the classified data:

$$n_{.j} = \sum_{i=1}^k n_{ij} \quad (1.6)$$

Primary accuracy measures are described below along with their equation (Liu *et al.*, 2007). First, the user accuracy (UA) for a given class i represents the number of correctly classified samples as belonging to this class and all samples classified in this class. Thus, UA also represents the incorrectly classified samples in class i while they should have been classified otherwise. For example, in the "Cropland" line of Table 1.1, over $n_{1.}$ that the system has classified as croplands, n_{12} are forests, n_{13} are grasslands, etc.

$$UA_i = \frac{n_{ii}}{n_{i.}} * 100 \quad (1.7)$$

The producer accuracy for a given class i shows the number of correctly classified samples and all samples belonging to class i with respect to the validation. It also represents the samples of the class i that are classified elsewhere than in class i . For example, in the "Grassland" column of Table 1.1, over $n_{.3}$ grasslands, n_{13} are classified as croplands, n_{23} are classified as grasslands, etc.

$$PA_i = \frac{n_{ii}}{n_{.i}} * 100 \quad (1.8)$$

The overall accuracy (OA) is the ratio of all the correctly classified samples over the total

validation samples (N):

$$OA = \frac{(100 * \sum_{k=1}^q n_{kk})}{N} \quad (1.9)$$

Finally the kappa (κ) coefficient gives a degree of agreement between the predicted and observed values:

$$\kappa = \frac{N_{..} \sum_{i=1}^k n_{ii} - \sum_{i=1}^k n_{i.} n_{.i}}{N_{..}^2 - \sum_{i=1}^k n_{i.} n_{.i}} \quad (1.10)$$

1.3.3.1 Image processing

There are four main steps when producing a land cover map. The first one is called: Image segmentation. Satellite imagery discretizes the land into pixels. There are two conceptual models conceived to characterize the spatial dimension: raster image made of pixels and vector image made of objects. Image segmentation is required when the pixel size is much smaller than land cover elements. However, when spatial resolution is close or a bit larger than the size of the land objects, image segmentation is not necessary because the land cover information is extracted at pixel scale. Then, by using remote sensing images or time series, the feature extraction step resides in computing the most discriminant variables to be inserted as input for the classification algorithm (Defourny, 2017a). There are different kinds of features:

- spectral features: multispectral reflectance, derived indexes, vegetation indexes;
- temporal features;
- textural features: local contrast, entropy, etc;
- spatial features (object-based approach).

The third step is the classification. It is composed of one or several numerical processes to assign each pixel or object to a class of land cover previously defined among a series of classes. A class of pixels or objects is characterized by a single vector which enable to recognize the pixels or objects from their vector. There are two types of classification. First, the supervised classification which consists in using training data in order to calibrate the algorithm. And then, the unsupervised classification for which there is no prior information about the pixels or objects to study. Nowadays the variety of methods to classify remotely sensed images in land cover classes is in constant expansion. Therefore, a few of these methods will be presented below.

Maximum Likelihood The Maximum Likelihood (ML) classification method was the most widely used classification method until very recently. In practice, the ML decision rule is based on probability. The Maximum Likelihood classifier assumes that the training

data for each class in each band are normally distributed, by targeting classes statistically using their multivariate probability density functions, and it calculates the probability that a given pixel belongs to a specific class. Each pixel is assigned to the class of which it has the highest probability (likelihood) to be a part of (Yang *et al.*, 2011); Defourny, 2017a). Even if the main strength of the ML classifier is the the total control of the analyst over the land cover classes that are employed in the final classification, there are several limitations to this method:

- the reliance on Gaussian distribution of input data;
- same set of features for all classes;
- high number of computations in order to complete the classification.

Due to these limitations, new methods have been developed with the purpose to overcome the ML classification method limitations.

Decision trees Decision trees (DTs) are image classification algorithms working like a "white" box, and are based on recursive binary partitions complying with a set of optimized rules (Defourny, 2017a). Features are the inputs in DTs which return an output, i.e. the decision, through a series of tests. Indeed, the whole tree is easy to interpret from its general structure to its ends, the process implementation is fast, data can be managed on different scales and the use of DTs allows feature selection (Peña-Barragan *et al.*, 2011).

Random Forest The Random Forest is a non-parametric machine learning algorithm becoming more and more used to identify crop types. The advantages of the Random Forest classifier are its stability and efficiency to obtain a high level of accuracy. Moreover, it can handle a large amount of spectral and temporal features (Hovis et al., 2017).

The strengths and weaknesses of the presented classification algorithm are regrouped in Table 1.2 (Defourny, 2017a).

Finally, postprocessing procedures are designed to enhance classification outputs by applying filtering operations or fusing several classification outputs. First, majority filters can remove macroscopic errors such as the salt-and-pepper effect produced by pixel-based classification. This kind of filters might also be applied to pixel-based classification output using objects obtained by multispectral reflectance image segmentation. Then, fusion techniques require the merger of the outputs.

Table 1.2: Strengths and weaknesses of algorithms used for classification of satellite imagery (Defourny, 2017a;Autre Gomez)

Algorithm	Strengths	Weaknesses
Maximum Likelihood	Simple application Easy to understand and interpret Predicts class membership probability	Parametric Assumes normal distribution of data Large training sample necessary
Decision trees	No need for any kind of parameter Easy to apply and interpret Handles missing data Handles data of different types (e.g. continuous, categorical) and scales Handles non-linear relationships Insensitive to noise	Sensitive to noise Tends to overfit Does not perform as well as others in large feature spaces Large training sample required
Random Forest	Capacity to determine variable importance Robust to data reduction Does not overfit Produces unbiased accuracy estimate Higher accuracy than DTs	Decision rules unknown (black box) Computationally intense Requires input parameters

1.3.4 Examples

In this subsection, three researches on crop identification will be presented as examples to illustrate the theory previously introduced in the previous subsections.

1.3.4.1 Crop identification using single-date multispectral imagery

The following researches used multi-date imagery for crop identification, generally during a given year or sometimes several years, because they acknowledged that the advantages of this method to complete the identification task. Nevertheless, single-date multispectral imagery can also be used for crop identification. In this context, Yang *et al.* (2011) carried out a research to evaluate the high resolution SPOT 5 satellite single-date imagery for crop identification using five supervised classification methods: minimum distance,

Mahalanobis distance, maximum likelihood, spectral angle mapper (SAM), and support vector machine (SVM).

- Minimum distance: this classifier uses the class means derived from the training data and assigns each pixel to the class that has the closest Euclidean distance from the pixel (Yang *et al.*, 2011);
- Mahalanobis distance: this method is similar to minimum distance, except that the covariance matrix is used in the calculation. Each pixel is assigned to the class for which Mahalanobis distance is the smallest (Yang *et al.*, 2011);
- Maximum likelihood: the maximum likelihood classification assumes that the data for each class in each band are normally distributed and it calculates the probability that a given pixel belongs to a specific class. Each pixel is assigned to the class that has the highest probability (Yang *et al.*, 2011);
- Spectral angle mapper (SAM): this classifier assigns pixels to the classes based on minimum spectral angles (Yang *et al.*, 2011);
- Support vector machine (SVM): the SVM classifier is a kernel-based machine learning technique that separates the classes with a decision surface that maximizes the margin between the classes (Yang *et al.*, 2011).

They chose two intensively cropped areas in Texas (United States) as study zones. In these areas, crop and non-crop classes were identified. The crop classes were: corn, cotton, grain sorghum and sugarcane; and the non-crop classes were: grass, mixed herbaceous species, mixed woody species, fallow, and water.

SPOT 5 provides 10-meter spatial resolution images composed of four spectral bands: green, red, near-infrared (NIR) and short wave infrared (SWIR). The researchers actually classified the image according to the five classification methods previously mentioned and they assessed the quality of each classifier in their survey.

Visual analysis of the classification maps produced reveals that crop and non-crop classes are well separated. However, even if fields in the maps were assigned one dominant class, some had small inclusions of other classes due to similar patterns among the classes according to the researchers. An accuracy assessment was performed for each classification methods and then, the statistics of each method was analyzed. Overall accuracy for the maximum likelihood classifier was the highest of the five methods used : 91% for the first study site and 87,4% for the second one. Moreover, maximum likelihood and support vector machine (SVM) were the best overall classifications among the five methods. As overall accuracy represents the overall performance of a classification method, producer

accuracy and user accuracy give a better insight into the accuracy of individual classes. "Producer's accuracy of a class indicates the probability of actual areas for that class being correctly classified, and user's accuracy of a class indicates the probability that areas classified as that class on the map actually represents the class on the ground" (Yang *et al.*, 2011). Once again, the best producer and user accuracies were provided by the maximum likelihood and SVM methods.

The same classification methods were tested on the four-band image for a 20-meter spatial resolution. After the analysis of the five classification methods accuracy assessment, the authors realized that these results were very similar to the 10-meter results. So, they concluded that a larger pixel size doesn't have an impact on the classification accuracy.

Finally, they wanted to assess if the number of spectral bands used in the methods had a significant impact on the classification accuracy, so they removed the SWIR band in this last test. The conclusions drawn from the removal of the SWIR band are that overall accuracy is higher for the four-band image and that the three-band classification maps are noisier. So, the SWIR spectral band might improve the classification results.

In conclusion, the maximum likelihood and the support vector machine classifiers were more effective than the others. Increasing the pixel size did not have an important effect on the classifications results. The addition of the SWIR band proved to be efficient. So, single-date imagery might be used in order to identify crops if the date is chosen depending on the optimum crop discrimination period.

1.3.4.2 Parcel-based crop identification

First, Schmedtmann *et al.* (2015) carried out a research to assess the use of remote sensing for crop identification in order to control the subsidy allowed by the Common Agricultural Policy (CAP) in Europe. As 40% of the European Union budget is allocated to CAP subsidies, it is important to ensure the appropriateness of these funds. This study provides a parcel-based crop identification on many parcels as possible to delete the usage of computer assisted photo-interpretation, which was the method previously used to control the parcels. However, this technique has a great disadvantage: it is rather long and it requires a team to analyze the photographs. The parcel-based crop identification is intended to produce an Automatic Control System more simple and cost-effective.

The study was implemented in a zone of 6390 km², in a Portuguese province northeast of Lisbon using multispectral Landsat 7 ETM¹¹ time series and data from the Portuguese

¹¹Enhanced Thematic Mapper

agency who retrieve the land use information for the European Union CAP. The data used was acquired by Landsat 7 from November 2004 to August 2005. Then, five optimal periods were identified according to their importance in the growth cycle of winter and summer crops in Portugal. To complete the identification, twelve crop classes were chosen: permanent grassland, forage crops, maize, rice, fallow, wheat, poor grassland, vineyard, non used area, barley, oat, and olive grove.

The Automatic Control System is composed of three steps. The first one is the classification which needs a prior variable selection in order to reduce the dimensionality in the dataset to prevent redundancy and over-fitting. Then, the classification step is completed using the Support Vector Machines classifier. The second step is the calibration and the final step is the application of the Automatic Control System.

The classification led to an overall accuracy of 68%, which is considered low by the authors and that is due to spectral similarities of the crop classes. To improve the overall accuracy, the reliability level was set to 80% and the overall accuracy reached 84% instead of 68%. However, this reliability level increase led to the decrease of the number of parcels automatically controlled. Another observation of the study led to the conclusion that some classes cannot be distinguished from one to another, so there would be no automatic classification for these classes: fallow, non used area and poor grassland, because they show the same land cover patterns.

Nevertheless, for some crop classes (rice, maize, wheat or vineyard), more than 80% of the studied parcels can be controlled by automatic crop identification instead of classic photo interpretation of remote sensing images. According to the authors, these classes were the easiest to identify.

Finally, controlling CAP area-based subsidies with remote sensing is essential for European authorities. In this survey, the researchers came up with an automatic classification system that combines automatic crop classification with reliability requirement. Setting up a reliability requirement allowed the authors to perform the classification even if it is complicated to distinguish some types of crops.

1.3.4.3 Object-based crop identification

Another methodology is proposed by Peña-Barragán *et al.* (2011), and it assembles object-based image analysis with decision trees in order to evaluate a large amount of crop types and vegetation status in a Californian county. A decision tree is an image classification algorithm working like a "white" box. Indeed, the whole tree is easy to interpret from its general structure to its ends. Decision trees provide various advantages such as data management on different scales, feature selection, fast processing, etc.

In this study, Peña-Barragán *et al.* (2011) wanted to study the links between various vegetation indexes and the studied crops at parcel scale. They also intended to assess the contribution of the different regions of the electromagnetic spectrum in crop identification. And finally, they tried to define how different growing-season dates influence crop discrimination. To carry out this survey, three key dates were chosen because they represented important growing periods: May 7th, June 24th and September 12th (2006). For these dates, images from the Terra Advanced Spaceborne Thermal Emission and Reflection Radiometer¹² were acquired. Spectral bands from ASTER were selected in the visible, near-infrared and shortwave infrared with 15 meters and 30 meters of spatial resolution. From these different wavebands, twenty vegetation indexes were extracted and classified according to the spectral regions used for the indexes computing.

The methodology developed by Peña-Barragán *et al.* (2011) is called: Object-based Crop Identification and Mapping (OCIM). It is divided into two sequential phases. First, the crop-fields were delimited by image segmentation with the purpose to apply, then, decision rules based on thresholds designed to classify every studied crop according to its spectral features, growth period and cycle, and its associated cultural practices. After the first phase, a sample of 1300 fields was chosen randomly over the study zone and the number of parcels for each crop was selected in accordance with the relative proportion of each crop over the total area. The object-based approach benefits from assessing new spectral, textural, contextual and hierarchical features compared to pixel-based identification methods.

The OCIM methodology was implemented using "the capacity of the decision tree approach to select the object-based features and their cutting thresholds that best fitted to every type of crop" (Peña-Barragán *et al.*, 2011).

Therefore, crop identification was divided into several parts according to crop calendar variability. The decision tree was built using NDVI thresholds depending on the growing-season period, the type of crop, and spectral and textural features (see Figure A.1, page 88). For example, in order to discriminate the general groups (i.e. winter cereals, meadows, summer crops and permanent crops), a NDVI threshold value was set to 0,25.

This survey uncovered that spectral and texture features used for crop identification in the OCIM methodology were affected by the spectral regions and the growing-season period. Vegetation indexes defined in the Near-InfraRed and red regions are very impor-

¹²ASTER

tant when it comes to observing the plant growth cycle. In this respect, the Normalized Difference Vegetation Index's (Section 1.3.2.1) contribution to the models was estimated to 50%.

As a matter of fact, the Short-Wave InfraRed region is also interesting to gain a better insight into plant development. However, the researchers concluded that the vegetation indexes characterized in the SWIR region had a higher influence in the model building (see Figure A.1, page 88) when they were combined rather than individually. This kind of vegetation indexes was also powerful in the crop identification scheme due to its sensitivity to crop variability.

Then, the researchers realized that texture features did not contribute to the model building in a significant way. Nevertheless, they helped in reducing the uncertainty in the model. (Peña-Barragán *et al.*, 2011).

Chapter 2

Methodology

2.1 Study zone

The study zone is located in central eastern Belgium across three provinces : "Brabant wallon", "Namur" and "Liège". It is also divided into eight municipalities : Jodoigne, Hélécine, Lincent, Orp-Jauche, Hannut, Wasseiges, Eghezée, Ramillies (Figure 2.1). It is a square of 12,5 kilometers sidewise, for a total of 15625 hectares, and it is also located in the agricultural region called: loam region.

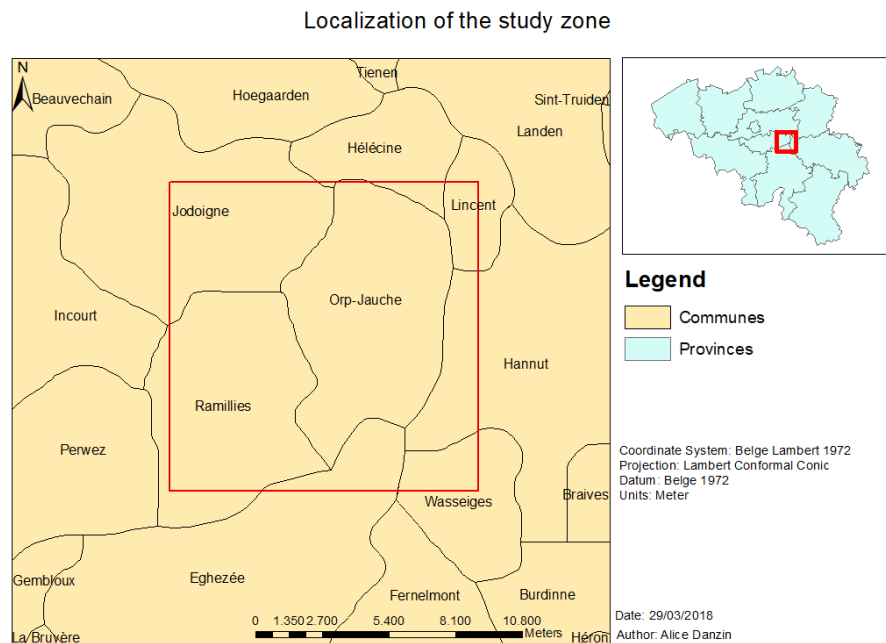


Figure 2.1: Location of the study zone over the different municipalities

The study zone includes 4300 agricultural plots. This zone is one of the most muddy flood affected areas in central Belgium (Janssens *et al.*, 2012) and it was also the study zone of Miguel Urdiaín master thesis, which explains its choice. Figure 2.2 provides the land use of the study zone displayed by the SPW. As a matter of fact, agricultural lands are the major land use type in the study zone. Agricultural land total surface represents 12907 hectares (82,1 % of the total study zone surface) which is much higher than in the other walloon municipalities where the agriculture land average reaches 53 %. Cities and artificialised territories account for 2200 hectares (14,1 %). Finally, forests and water bodies represent 452 hectares (2,9 %) and 66 ha (0,4 %), respectively.

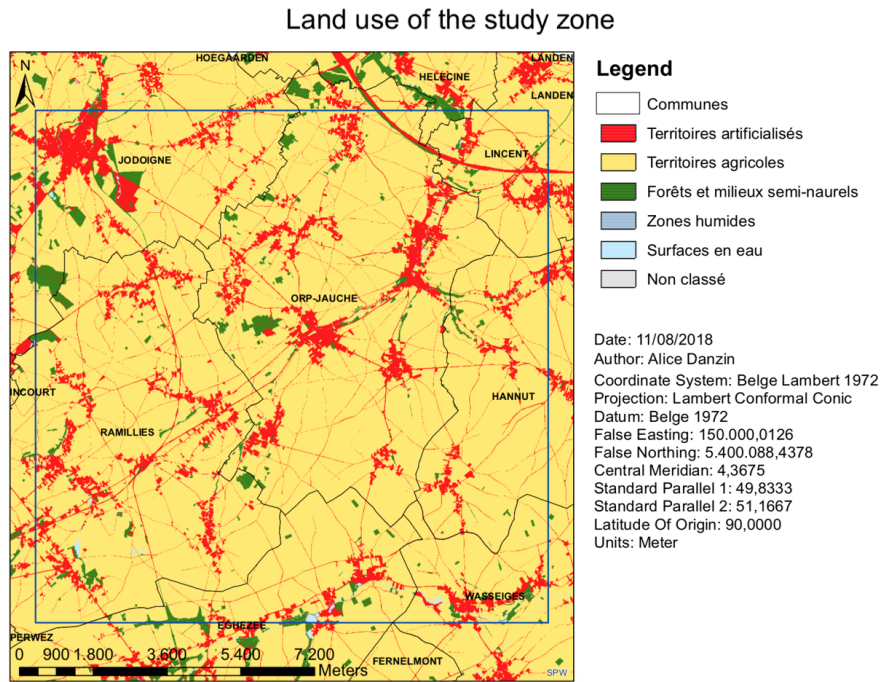


Figure 2.2: Land use of the study zone, delineated by the blue square (Data source: SPW)

2.2 Data

This section refers to the overall data collected and used to meet the main objective of this master thesis.

2.2.1 Land Parcel Identification System map

Every year, the SPW publishes an anonymous cadastral map for agricultural land based on the land area declarations filled by the Belgian farmers to specify which crop they plant on each agricultural plot. The map localizes the extent of each parcel using polygons and

refers to the main crop cultivated on each plot. So, cover crops aren't specified on the anonymous cadastral map. In this master thesis, the LPIS maps used were the ones corresponding to the standing crops in 2016 and 2017.

2.2.2 Sentinel-2 products

Several products from the Sentinel-2 satellites are used in this master thesis as well as products from the Sentinel-2 for Agriculture program.

Sentinel-2 provides different kinds of products. The first one used in this master thesis is the Level-2A product which provides bottom of atmosphere reflectance images. The L2A products are basically 100x100 km² tiles in cartographic geometry using the WGS84 UTM 31N in this case.

Then, the L3B product is composed of vegetation status indicators, i.e. vegetation indexes. The biophysical indicators used in this case are the Normalized Difference Vegetation Index (NDVI) and the Leaf Area Index (LAI), describing the vegetative development of the crops on an acquisition basis.

Each LAI and NDVI image is directly derived from Sentinel-2. The pixel value is the LAI multiplied by 1000. The spatial resolution is 10 meters. There is one spectral band and the data format is GEOTIFF. The volume of each image is approximately 400 Mo.

2.3 Softwares

QGIS QuantumGIS (2.18.15 version) is an open source software and a professional Geographical Information System application. A GIS is the techniques and methods of acquisition, organization, analysis and edition of data spatially referenced. QGIS is mainly used to visualize, edit, analyze data and compose printable maps (QGIS).

Arcmap Arcmap (10.5.1 version) is a software that allows to analyze and use Geographical Information Systems (GIS). This software allows the user to create maps, analyze geographic data and use toolboxes in order to build small models, for example.

Matlab Matlab is a programming language derived from multi-paradigma environment numerical computing called also Matlab.

2.4 Overall structure of the methodology

Figure 2.3 represents the overall structure of the different steps developed in the methodology of this master thesis. Then, each step will be described in details in the following sections.

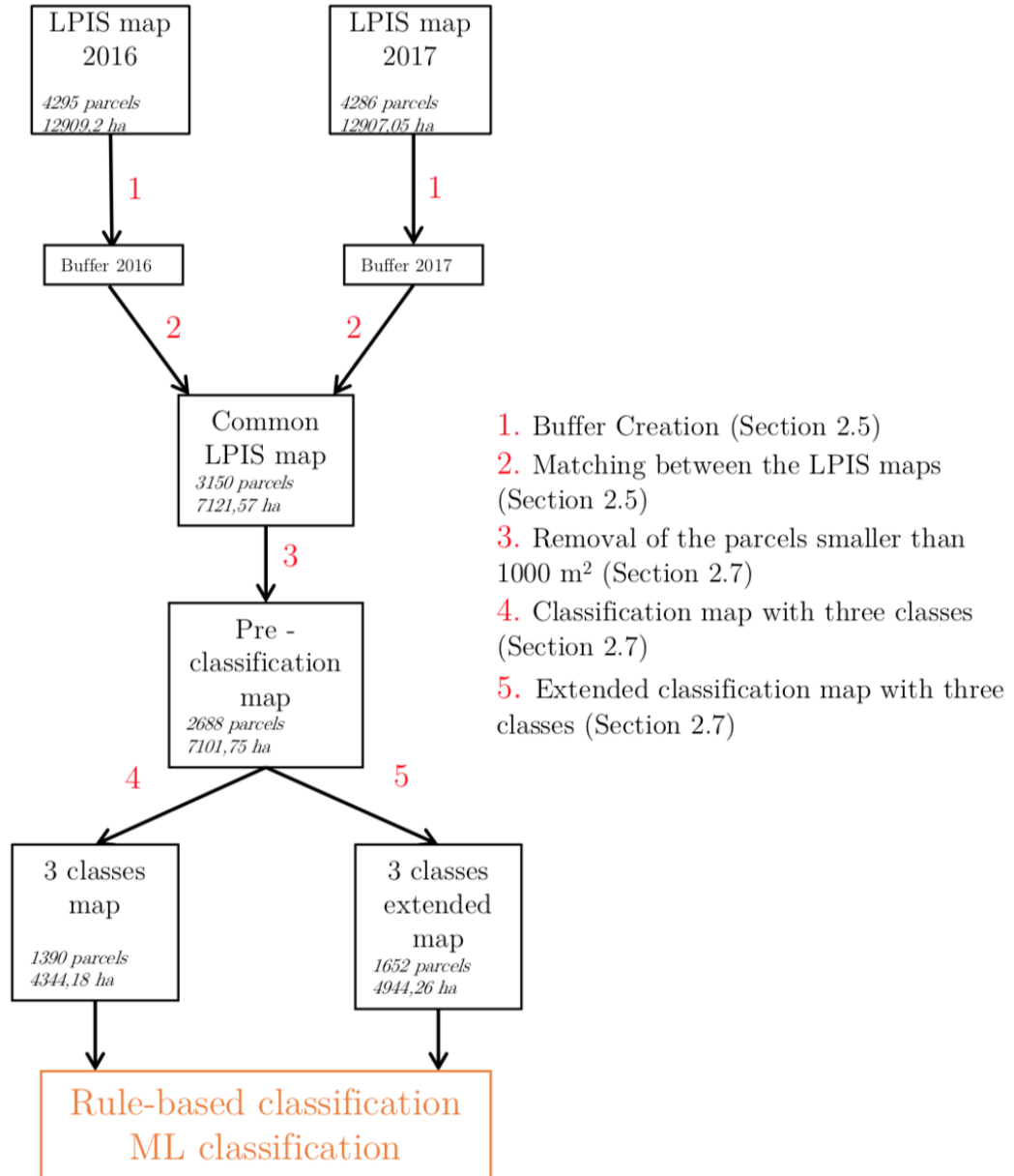


Figure 2.3: Diagram representing the different steps of the transformation of the LPIS maps into the classification maps

2.5 Image processing

The data used in this master thesis was collected from March 2016 to April 2018. As it will be explained later on, all methodology steps were performed mainly performed on the 2016 and 2017 data. The first step in the data treatment is to create a buffer of the 2016 and 2017 LPIS maps. This was achieved with the *Arcmap* software and using the **Buffer** tool. In order to eliminate the mixels, i.e. 1,5 pixels, 15 meters were removed from each shapefile because the spatial resolution of the Sentinel-2 products (NDVI images, 10 band images, etc) is 10 meters.

Then, the LAI and NDVI images had to be fitted with the LPIS shapefiles (maps). This operation is possible with a *QGIS* python plugin : **Clipper**. This plugin uses a plugin function in the same shapefile selecting a line or polygon as clipping feature and then clips all overlaying features. For this step, *QGIS* was chosen over *Arcmap* because the process is faster on *QGIS*.

Once each image is fitted to the LPIS maps, the statistics of these images can be calculated. Indeed, each plot of land in the study zone is composed of a various number of pixels (10x10 meters). So, in order to acknowledge the statistics of LAI and NDVI values corresponding to the crops in place on each plot of land. The statistics of the images values were retrieved for all parcels in the study zone and for each date available using the Spatial Analyst tool : **Zonal statistic as a table**, in *Arcmap* and the **Zonal Statistics Plugin** in *QGIS*. Next, still using the *Arcmap* software, the statistic tables obtained in the previous step were assembled with the LPIS shapefiles tables in order to match the statistics with the corresponding crop plots. This matching was carried out using the *Arcmap* function: **Joins and relates**.

After that, all the joined tables were saved in a format readable in *Excel* (.dbf). Analyzing the tables in *Excel* requires using the function **VLOOKUP** and the use of a pivot table in order to match the identification number of the parcels (referenced in the LPIS maps) from one year to another. At the end of this process, the common LPIS map for 2016 and 2017 contains 3150 parcels and has a surface of 7121,57 hectares (Figure 2.3). This represents 55,2% of the surface of the original LPIS maps. It has dropped to this value due to the implementation of the buffer, as explained previously, and also because some 472 parcels were not cultivated in 2017 (no connection with 2016). In the further methodology steps of this master thesis, LAI images won't be used because the NDVI is less affected by errors due to atmospheric variations, for example, than the LAI.

2.6 Crop rotations analysis based on the NDVI time series

Different scenarios will be analyzed in order to make decision rules to achieve the intended predictions. To get there, Biielders *et al.* (2011) have identified the main rotations in the different agricultural regions of Belgium (Table 2.1).

Table 2.1: Main crop rotations identified of the different agricultural regions in Belgium by the GISER Convention (Biielders et al., 2011 [4])

Crop rotation patterns	Crop rotations
Monoculture	Maize Winter Wheat Winter Barley
Succession of two crops	Winter Wheat-Sugarbeet Winter Wheat-Maize Winter Wheat-Potato Winter Wheat-Linen Winter Wheat-Rapeseed Winter Wheat-Winter Barley Winter Barley-Potato Winter Barley-Linen Winter Barley-Rapeseed Spelt-Sugarbeet Spelt-Maize Spelt-Potato
Succession of three crops	Sugarbeet-Winter Wheat-Potato Sugarbeet-Winter Wheat-Maize Sugarbeet-Winter Wheat-Linen Maize-Winter Wheat-Potato Potato-Winter Barley-Sugarbeet Linen-Winter Barley-Sugarbeet Potato-Winter Barley-Maize Winter Barley-Potato-Winter Wheat Winter Barley-Linen-Winter Wheat Winter Barley-Rapeseed-Winter Wheat Winter Wheat-Winter Barley-Maize Winter Wheat-Winter Barley-Sugarbeet

Thanks to the analysis of the NDVI values of the main crops (sugarbeet, maize, potato, winter wheat, winter barley), schematic NDVI profiles for each main crop could be built as well as schematic profiles for the traditional winter cover crops in central Belgium (Figure B.1 in Appendix B).

2.7 LPIS map processing

The LPIS shapefile of 2017 was used as the reference map for the different classification methods. In section 2.5, a buffer was applied to the LPIS shapefile in order to eliminate border pixels. This resulted in a reduction of the parcel size. In order to remove small parcels from the LPIS shapefile, the parcel surface was estimated thanks to the **Calculate Geometry** function in a new field created for this aim. As the pixel size is 10x10 meters, the surface pixel equals to 100 m². So, a threshold of 1000 m² surface was set to remove the parcels whose surface is below that threshold in the shapefile. This new shapefile is called the pre-classification map (Figure 2.3), it contains 2688 parcels and has a surface of 7101,75 hectares.

To determine whether winter cereals and winter cover crops can be distinguished or not, three classes of interest were selected in the reference map (i.e. LPIS shapefile): winter wheat, winter barley and the spring crops were assembled in one class: winter cover crops (WCC). The WCC class includes all the parcels cultivated with maize, sugarbeet, potato in 2017 and that were preceded by winter wheat or winter barley in 2016.

Each class was created by selecting its attributes among the entire LPIS attribute table and then by exporting the selected data into a new shapefile. After, the three newly created shapefiles were assembled using the **MERGE** tool in order to complete the classification map. In this map, two new fields were added: "CLASS" and "Value" (see Table 2.2), with the purpose of using these fields as reference fields for the classification methods (see following sections). This classification map is called the "3 classes map", it is composed of 1390 parcels and has a surface of 4344,18 hectares. It should be, furthermore, noted that the bare soils are included in each class depending on the time of year.

Table 2.2: Data assigned to the new fields "Value" and "CLASS"

Value	CLASS
1	Winter wheat
2	Winter barley
3	Winter Cover Crops (WCC)

An extended classification map was created including all antecedents of maize, sugarbeet, potato. This map will also be used in the rule-based classification and Maximum Likelihood classification (see Sections 3.2 and 3.3). The map is called: 3 classes extended map, it is composed of 1652 parcels and has a surface of 4944,26 hectares.

2.8 Rule-based Classification

In this section, the first classification method will be presented. This first method is a rule-based classification procedure. It is a thresholding classification method. The principle of this technique is to identify the minimum and maximum spectral values in a spectral band for each crop class. In this master thesis, the rule-based classification method uses defined NDVI thresholds with the purpose of classifying the study zone parcels according to three classes: "Winter wheat" (1), "Winter barley" (2) and "Winter Cover Crops" (3).

The NDVI thresholds corresponding to each class will be determined in Section 3.2. In order to determine these thresholds, an analysis of the NDVI time series between September 2016 and March 2017 of the three classes was performed. It was achieved by using the common LPIS map processed in Section 2.5 with the corresponding NDVI time series. This also allowed to observe the crop rotation patterns on any given plot from 2016 to 2017. Then, plots with winter wheat and winter barley in 2017 were filtered. Regarding winter cover crops, parcels with sugarbeet, maize and potato in 2017 were assembled when they were preceded by a winter cereal in 2016 (winter wheat or winter barley). As a matter of fact, a winter cover crop is mostly implemented between a winter cereal and a spring crop. In order to complete the analysis, NDVI averages and standard deviations were calculated for each date in Excel.

First, a test was carried out using four classes. Indeed, the class "bare soil" was added to the three others. This new class is set to contain the NDVI values between 0 and 0,19. Then, a second test was set up with only the three classes mentioned previously: winter wheat, winter barley and winter cover crops. Both tests were carried out for a pixel-based approach and a parcel-based approach. In the pixel-based approach, each value of each pixel is used in the classification. However, in the parcel-based approach, the pixel average is computed at parcel scale, using the **Zonal Statistics** tool in Arcmap, so that each pixel in each parcel has the same value as the other pixels of the parcel.

Once the NDVI thresholds were identified, they were applied to the corresponding NDVI image using the Arcmap tool **Reclassify**. This tool allows to determine the number of classes of interest to insert in a raster, i.e. the NDVI image. It also allows to define the value range of the classes manually. In order to assess the quality of the rule-based classification, confusion matrices were calculated using the **Semi-Automatic Classification Plugin** in QGIS. The results obtained from these confusion matrices will be presented in Section 3.2.

2.9 Maximum Likelihood Classification

Many researchers have used the Maximum Likelihood classification method in crop identification (Yang *et al.*, 2011). In order to determine if winter cereals can be distinguished from winter cover crops in autumn and winter, this method will therefore be used.

2.9.1 Pre-classification steps

Before starting the classification itself, some modifications need to be added to the reference classification map obtained in Section 2.7. This shapefile was divided into two other shapefiles in order to obtain training areas for calibration and a shapefile to validate the classification. To achieve this purpose, a field was added in the attribute table of the LPIS shapefile using the function **Add field**. In this field, a random value between 0 and 1 was assigned to every parcel using the **Field Calculator** function. Then, the calibration and validation maps were created using the **Select by Attribute** function by choosing all values under 0,5 for calibration and the rest for validation.

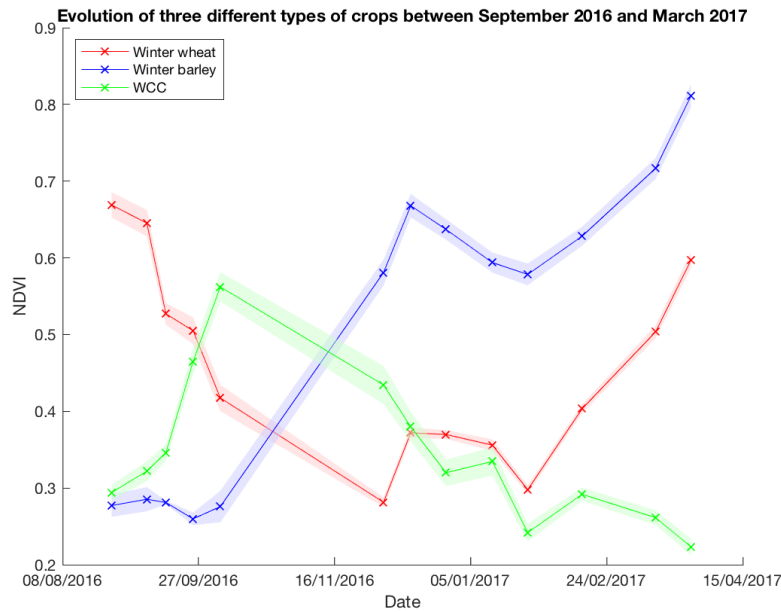


Figure 2.4: Graphical NDVI evolution of winter wheat, winter barley and winter cover crops (regrouped in one class) between September 2016 and March 2017. The color bands represent the 95% confidence intervals of the mean values

Then, with the help of the data retrieved from the NDVI images (see section 2.5), general tendencies for winter wheat, winter barley and winter cover crops were built. The solid line in Figure 2.4 represents the NDVI mean of the sample data for each type of

crops. Then, the coloured patch around each solid line delimits the confidence interval on the NDVI mean of each crop. The NDVI means and confidence intervals were calculated using the Excel functions **AVERAGE** and **CONFIDENCE.NORM** with a confidence level of 95%, respectively.

In Figure 2.4, winter barley is represented in blue. This crop is generally sown in October (as previously mentioned). So, its growth starts earlier than the winter wheat growth. However, a NDVI average of 0.8 in March seems rather high at first. But, in reality, winter barley is at the end of the tillering stage at this time of the year and the crop covers almost all the soil surface as shown in Figure 2.5 (Bertin, 2017). Furthermore, the analysis of the 14th of March 2017 NDVI image shows that the majority of the winter barley parcels in the study zone present NDVI values higher than 0.75.

Then, as winter wheat (red) can be sown from October until November, its growth will start later than winter barley. This is what will allow to distinguish these two winter cereals. And finally, the green line was built by selecting the winter cereals in 2016 followed by spring crops (sugarbeet, maize and potato) in 2017, in the NDVI time series. This scenario is the most suitable to the implementation of a winter cover crop between these two kinds of crops. The green line in Figure 2.4 clearly shows a NDVI average of 0.6 between September and December followed by a decrease of the NDVI average to 0.2 in winter which corresponds to the signature of winter cover crops.

Thanks to Figure 2.4, temporal windows could be identified in order to select the images used for the classification:

- 25th September 2016,
- 5th October 2016,
- 4th December 2016,
- 27th December 2016,
- 15th February 2017,
- 14th March 2017.



Figure 2.5: Winter wheat at the end of the tillering stage (Bertin, 2017)

Once visualizing the NDVI image for the 5th of October 2016, it was removed from the temporal windows used in the Maximum Likelihood classification because it is filled with clouds and therefore unusable for the classification.

The Maximum Likelihood classification will be performed using several type of images:

- the NDVI images;
- the 10 band images ("images brutes"), L2A products;
- the Soil Adjusted Vegetation Index, calculated from the 10 bands images;
- the Normalized Difference Water Index, calculated from the 10 bands images;
- the Brightness Index, also calculated from the 10 bands images.

Then, the reference classification map created in Section 2.7 was used to clip all the different types of images previously mentioned that will be used for the maximum likelihood classification. This step was done in *QGIS* using the raster tool: **Clipper**.

2.9.2 Maximum Likelihood classification methodology

Once the images were clipped, the actual maximum likelihood classification can be done using both Arcmap and QGIS. To achieve this purpose, the first geoprocessing tool used is **Create Signatures** (Arcmap). It creates an ASCII signature file of classes defined by input sample data which is the calibration shapefile derived from the LPSI map (see Section 2.7), and a set of raster bands, i.e. the different types of images defined in the

previous subsection. The output signature file has the .gsg extension required for the following step.

Next, thanks to the signature file, the **Maximum Likelihood Classification** can be performed using the L2A and L3B products, and the signature file as inputs specifying that all classes will have the same a priori probability. The output is a classified raster which was then transformed into a TIFF file in order to perform the accuracy assessment in QGIS. The Maximum Classification Methodology is set up using a pixel-based approach. This methodology was carried out using different vegetation indices, different type of images and different dates. It will be explained in Section 3.3 and the results of this classifications will also be presented in that section.

2.9.3 Accuracy assessment

In order to assess the quality of the classification, an accuracy assessment of each classified raster must be performed. This is carried out using the accuracy post-processing tool of the **Semi-Automatic Classification Plugin** in QGIS. The "classification to assess" corresponds to each classified raster generated in Arcmap (see previous subsection) and the "reference shapefile or raster" is the validation map created in Section 2.9.1. As the validation map is a shapefile, it needs to be turned into a raster using the Arcmap tool **Polygon to Raster**, and then into a TIFF file by exporting the raster created. This accuracy assessment generates confusion matrices that evaluate the accuracy of the classification performed for each identified scenario. The results of the accuracy assessment will be presented in the following chapter (in Section 3.3).

Chapter 3

Results and discussion

3.1 Crop rotation analysis based on the NDVI time series

In this section, first, basic statistics concerning the crops grown in the study zone in 2017 will be presented. Then, crop rotations in 2016 and 2017 will be analyzed thanks to the NDVI time series originated from the NDVI images retrieved from Sentinel-2 (see Section 2.5 for the processing details). The crop rotation analysis will be divided into five parts corresponding to the five main crops identified: sugarbeet, maize, potato, winter wheat and winter barley.

3.1.1 Basic statistics in 2017

As previously stated, in order to be able to work on the same parcels every year, some *Excel* manipulations were performed in Section 2.5 and the total number of parcels used in the time series is 3150 over a surface of 7121,57 hectares (it is called the common LPIS map). This surface represents 55,2% of the total agricultural surface delimited by the LPIS maps provided by the SPW (Figure 2.3). These 3150 parcels are those used to compute the basic statistics presented in this subsection.

First, Table 3.1 shows for each category of crops the number of parcels included in the category and the percentage in comparison with the total amount of parcels. In this table, four categories were determined. The first one is the spring crops and it accounts for all the spring crops: bean, carrot, chicory, leek, linen, maize, onion, pea, potato, spring oat, sugarbeet (Bertin, 2017). Then, the winter cereals category accounts for winter wheat and winter barley. The pastures category includes temporary and permanent pastures. And finally, the "Other crops" category regroups all the minor crops, i.e. crops that are not predominant in the sample of parcels.

Table 3.1: Basic statistics calculated for 2017 including the type of crops, the number of parcels the percentage of the type of crop compared to the total amount of parcels and the area in hectares

	Number of parcels	%	Area [ha]
Spring crops	1190	37,8	3377,95
Winter cereals	985	31,2	2874,93
Pastures (temporary and permanent)	685	21,7	464,51
Other crops	289	9,3	404,18
Total	3150	100	7121,57

In order to determine the importance of maize, sugarbeet and potato in the "Spring crops" category, Table 3.2 displays the number of parcels that accounts for the three crops just mentioned, as well as their proportion in comparison with the total number of spring crops and the total amount of parcels. The table also shows the surface occupied by these spring crops. It represent 29,7% of the common LPIS map surface (Section 2.5) used in these analyses, i.e. 7121,57 hectares (Table 3.1).

Table 3.2: Basic statistics calculated for 2017 including the type of spring crops, the number of parcels, the percentage of the type of crop compared to the total amount of spring crops parcels ($\%_{sp}$), the percentage of the type of crop compared to the total amount of parcels ($\%_{tot}$) and the area in hectares

	Number of parcels	$\%_{sp}$	$\%_{tot}$	Area [ha]
Sugarbeet	309	26	9,8	1023,06
Maize	287	24,1	9,1	216,12
Potato	268	22,5	8,5	876,72
Total	864	72,6	27,4	2115,9

Table 3.3 shows number of parcels of winter wheat and winter barley in 2017, the proportion of each crop in comparison with the total number of winter cereals and in comparison with the total number of parcels. Table 3.3 also displays the surface occupied by each crop. Winter wheat is obviously the most important among both winter cereals. Winter cereals represent 40,37% of the common LPIS map surface (Section 2.5) used in these analyses, i.e. 7121,57 hectares (Table 3.1).

Table 3.3: Basic statistics calculated for 2017 including the type of winter cereals, the number of parcels, the percentage of the type of crop compared to the total amount of winter cereals parcels ($\%_{wc}$), the percentage of the type of crop compared to the total amount of parcels ($\%_{tot}$) and the area in hectares

	Number of parcels	$\%_{wc}$	$\%_{tot}$	Area [ha]
Winter wheat	853	86,6	27	2566,33
Winter barley	132	14,4	4,2	308,6
Total	985	100	31,2	2874,93

Moreover, Table 3.4 illustrates the possible antecedents when there was sugarbeet, maize and potato in 2017. The antecedents comprise the predominant crops that preceded sugarbeet, maize and potato. The group "Others" include the minority crops. Table 3.4 also shows for each antecedent the number of parcels of this antecedent and its percentage in comparison with the number of parcels of the corresponding spring crop (sugarbeet, maize or potato).

Table 3.4: Basic statistics calculated for 2017 including the three main spring crops in 2017, what preceded each of them in 2016, the number of parcels and the proportion of each antecedent compared to the total number of parcels of the corresponding spring crop (i.e. maize, sugarbeet, potato)

2017	2016	Number of parcels	%
Sugarbeet	Winter wheat	162	52,4
	Winter barley	77	25
	Peas	21	6,8
	Maize	14	4,5
	Potato	10	3,2
	Others	25	8,1
Maize	Winter wheat	102	35,5
	Maize	90	31,3
	Winter barley	26	9
	Sugarbeet	15	5,2
	Potato	9	3,2
	Others	45	15,7
Potato	Winter wheat	165	61,5
	Winter barley	21	7,8
	Maize	20	7,5
	Sugarbeet	19	7
	Peas	15	5,6
	Others	28	10,5

Finally, Table 3.5 displays the possible types of crops in 2016 which preceded winter wheat and winter barley (grown in 2017). First, the antecedents of winter wheat were, mostly, potato, sugarbeet, maize, fiber flax, winter wheat, peas and chicory. There were also minority crops regrouped in the "Others" category which didn't exceed 10 parcels: fodders, open air vegetables, spring oat, carrot, fiber hemp, rapeseed, endive, faba bean, green beans, oilseed flax, winter barley, spring barley, temporary pasture and winter triticale. Then, the lower part of Table 3.5 shows the possible crops in 2016 preceding winter barley. The main antecedent of winter barley is winter wheat (82,6 %).

Table 3.5: Basic statistics calculated for 2017 for winter wheat and winter barley in 2017, what preceded each of them in 2016, the number of parcels, the proportion of each antecedent compared to the total number of parcels of the corresponding winter cereal

2017	2016	Number of parcels	%
Winter wheat	Potato	238	28
	Sugarbeet	206	24,3
	Maize	120	14,1
	Fiber flax	68	8
	Winter wheat	55	6,4
	Peas	47	5,5
	Chicory	40	4,7
	Others	77	9
Winter barley	Winter wheat	109	82,6
	Winter barley	7	5,3
	Fiber flax	7	5,3
	Spelt	5	3,8
	Rapeseed	1	0,75
	Green beans	1	0,75
	Maize	1	0,75
	Peas	1	0,75

3.1.2 Sugarbeet

Figure 3.1 represents the evolution of the NDVI average of sugarbeet crops in 2017 and their antecedents in 2016. In each case, the patch is delineated by the 97,5% and 2,5% percentiles around the mean for every category. The main antecedents are displayed on the graph: winter wheat in red, winter barley in blue and spring crops in green (maize, potato and sugarbeet), as identified in Table 3.4. The remaining antecedents are then regrouped in the "Others" category and includes: chicory, winter rapeseed, spelt, faba beans, green beans, fibre flax, peas, permanent pasture, temporary pasture (black).

The right part of Figure 3.1 represents the NDVI average time series of sugarbeet in 2017, starting from April 2017 when NDVI average rises after after seeding. Then the sugarbeet crops highly grow from April-May to October-November, which corresponds to the time of year when the sugarbeet harvest starts.

The major antecedent of sugarbeet crops in 2016 was winter wheat (Table 3.4) in red in Figure 3.1. The figure shows that between April and June 2016, the NDVI average is rather high. But then, it starts decreasing from June to September when winter wheat is

harvested. As a matter of fact, NDVI values drop throughout grain filling, i.e. when winter wheat color changes from green to yellow. After that, the NDVI average rises again to reach a maximum of 0,6 in October and then decreasing until December-January showing the evidence of the implementation of winter cover crops between winter wheat and sugarbeet.

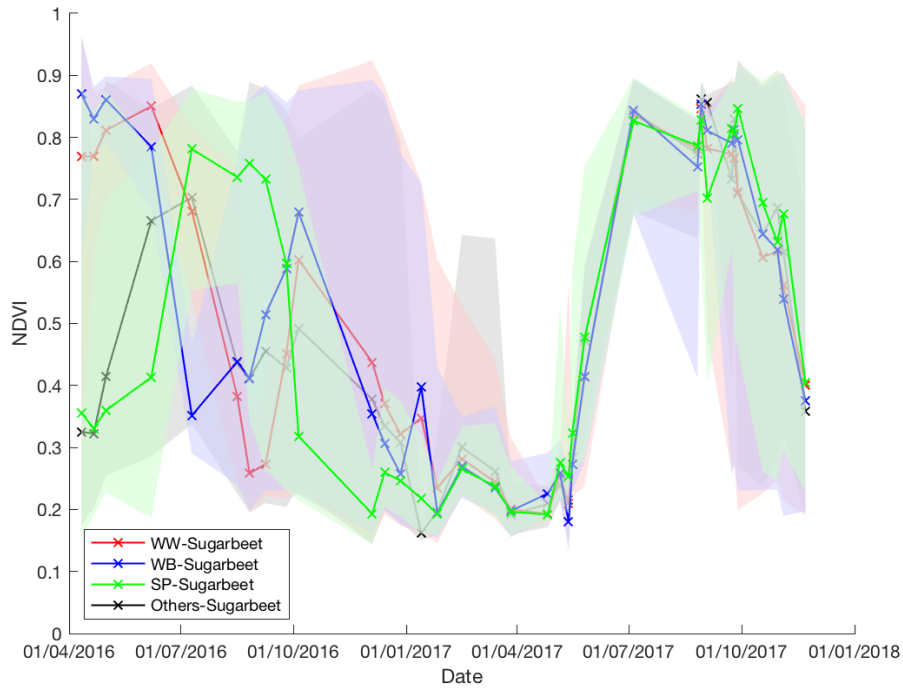


Figure 3.1: Graphical evolution of the NDVI average of sugarbeet crops in 2017 preceded by winter wheat crops (red), winter barley crops (blue), spring crops (green) or other crops (black) in 2016, showing the 97,5% and 2,5% percentiles around the mean as the coloured patches for each curve

Winter barley (in blue in Figure 3.1) is also a preponderant antecedent of sugarbeet crops (Table 3.4). In this case, NDVI average is also high from April until June. Then, it decreases in July when winter barley is harvested. This NDVI decrease corresponds to grain filling. Figure 3.1 shows that after July, NDVI values rises again from September to December clearly demonstrating the presence of winter cover crops between winter barley and sugarbeet crops.

When sugarbeet was preceded by maize, sugarbeet or potato (in green in Figure 3.1), the same trend as described previously in this subsection is seen. NDVI values increase from April to October and start dropping from October when these kinds of spring crops are harvested. The intercrop period between these spring crops and sugarbeet does not show

any sign of winter cover crops because weather conditions (i.e. cold temperatures that do not allow germination and plant growth, a soil too moist to till the land, ...) do not allow the implementation of winter cover crops during winter season.

3.1.3 Maize

Figure 3.2 shows the evolution of the NDVI average of maize crops in 2017 and their antecedents in 2016. Maize is mostly preceded by winter wheat in red, winter barley in blue and spring crops in green (maize, potato and sugarbeet). The "Others" category (in black) consists in the minority crops remaining: fodder, spring oat, winter rapeseed, spelt, faba beans, fiber flax, alfalfa, tree nurseries, permanent pasture and temporary pasture.

Maize crops are generally sown in March or in April, and start growing in May (right part of Figure 3.2). The maize growth continues until October. Then, the NDVI drops in November when it is harvested. The NDVI drop may also correspond to the moment becomes yellow in the case of grain maize. If it is green maize, the farmer does not wait until senescence to harvest.

In the study zone, winter wheat (in red in Figure 3.2) is the major antecedent of maize crops in 2016. NDVI average starts increasing in February when the winter wheat growth is triggered and it decreases at the beginning of July when the grain filling process starts. The winter wheat harvest then takes place in September. Figure 3.2 then shows that NDVI values rise again to drop further in December. This corresponds to the implementation of winter cover crops between winter wheat and maize crops.

Unexpectedly, maize is the second major antecedent for maize crops in 2016 (Table 3.4). It is combined with sugarbeet and potato in green in Figure 3.2. In the same way as maize in 2017, NDVI values increase from April to October and start dropping from October when these kinds of spring crops are harvested. The intercrop period between these spring crops and sugarbeet clearly shows sign of bare soils because weather conditions do not allow the implementation of winter cover crops during winter season.

Then, winter barley (in blue in Figure 3.2) is also a preponderant antecedent of maize crops (Table 3.4). In this case, NDVI average is also high from April until June. Then, it decreases in July when winter barley is harvested. This NDVI decrease corresponds to grain filling. Figure 3.2 shows that after July, NDVI values rise again from September to December clearly demonstrating the presence of winter cover crops between winter barley and maize crops.

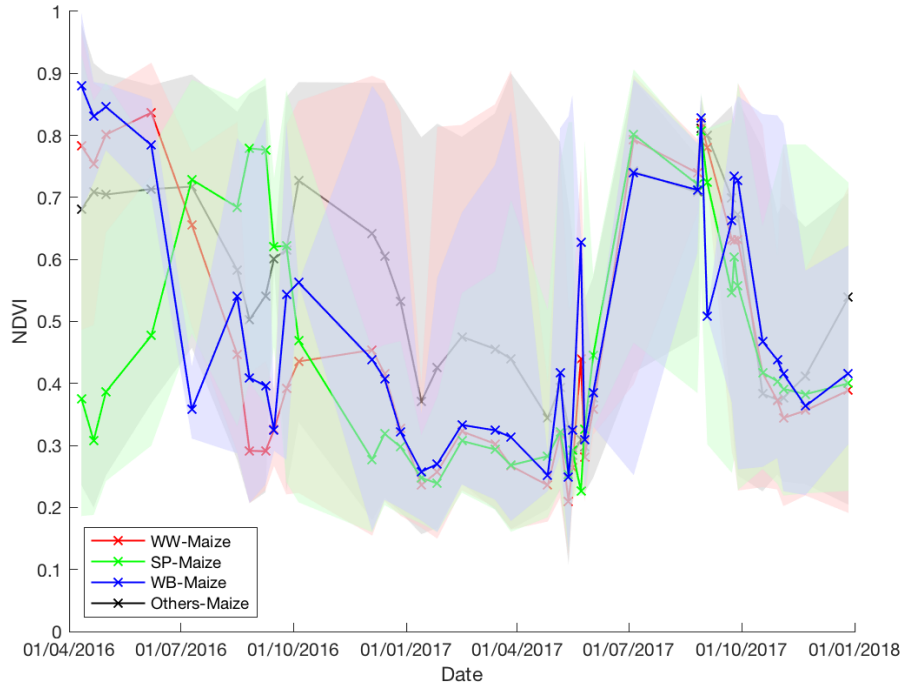


Figure 3.2: Graphical evolution of the NDVI average of maize crops in 2017 preceded by winter wheat crops (red), winter barley crops (blue), spring crops (green) or other crops (black) in 2016, showing the 97,5% and 2,5% percentiles around the mean as the coloured patches for each curve

3.1.4 Potato

Figure 3.3 displays the evolution of NDVI average of potato crops in 2017 and their antecedents in 2016. In this study zone, the major crops preceding potato: winter wheat in red, winter barley in blue and spring crops in green (maize and sugarbeet). Moreover, the "others" category (in black) regroups the minority crops: fodder, other vegetables, textile hemp, chicory, spelt, fiber flax, oilseed flax, tree nurseries and peas.

First of all, the evolution of the NDVI average of potato crops is representative of the reality. In fact, potatoes are sown in March. They start growing in May-June and are harvested in September.

The two major antecedents of potato crops in 2016 were winter wheat and winter barley (Table 3.4) just like for the sugarbeet crops (Subsection 3.1.2). The trends followed by these two winter cereals is the same as when they preceded sugarbeet crops, as it is explained in Subsection 3.1.2.

Then, potato crops were also preceded by maize and sugarbeet combined in green (Figure 3.3) in 2016. NDVI values increase from April to October and start dropping from October when these kinds of spring crops are harvested. The intercrop period between these spring crops and sugarbeet clearly shows sign of bare soils because weather conditions do not allow the implementation of winter cover crops during winter season.

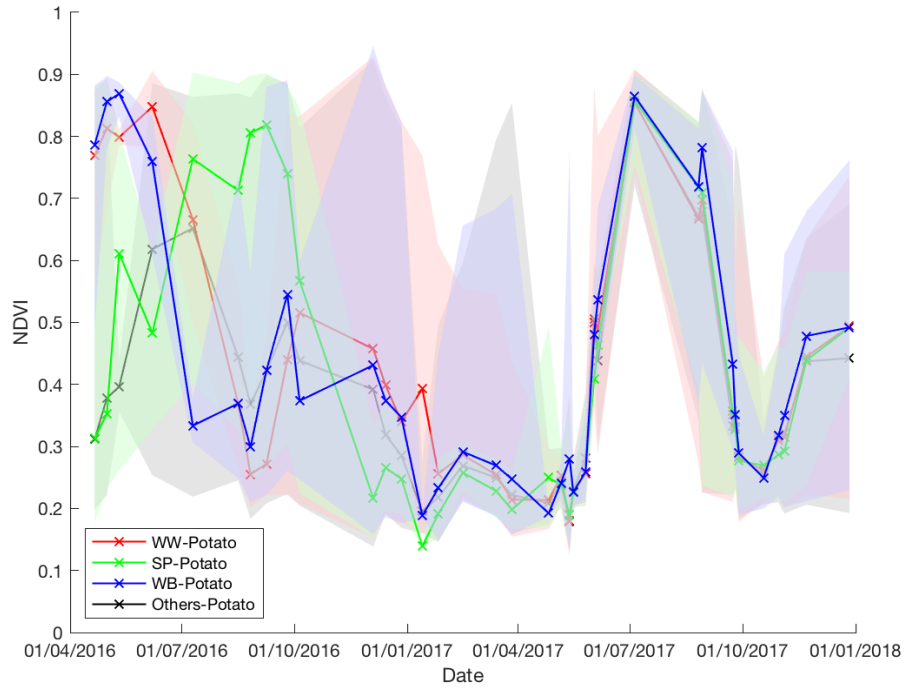


Figure 3.3: Graphical evolution of the NDVI average of potato crops in 2017 preceded by winter wheat crops (red), winter barley crops (blue), spring crops (green) or other crops (black) in 2016, showing the 97,5% and 2,5% percentiles around the mean as the coloured patches for each curve

3.1.5 Winter wheat

Figure 3.4 represents the evolution of winter wheat crops in 2017 and their antecedents in 2016. First, the crops preceding winter wheat in 2016 are: spring crops in green (maize, potato, sugarbeet) and winter cereals in magenta (winter wheat and winter barley). Then, the minority crops preceding winter wheat preceding are: fodder, other vegetable raised in open air, spring oat, carrot, chicory, winter rapeseed, endives, faba beans, green beans, oilseed flax, fiber flax, spring barley, peas, temporary pasture and winter triticale. Winter wheat cultivated in 2017 planted October-November can be harvested until the end of September, as seen in the right part of Figure 3.4.

The three major antecedent of winter wheat in 2016 (Table 3.5) were potato, sugar-beet and maize (in green in Figure 3.4). These crops start developing in April and NDVI average rises until September-October when potatoes can be harvested, maize and sugar-beet are harvested from October.

Then, winter cereals (winter wheat and winter barley) preceding winter wheat are represented in magenta in Figure 3.4. winter wheat is more often an antecedent of winter wheat than winter barley (Table 3.5). In the case of winter barley, the harvested in July. Then, if winter wheat is follows winter wheat, the first one is harvested in September and the second is planted in November (Figure 3.4). For these two cases, there is no possibility of implementing a winter cover crop between these kinds of winter cereals due to the sort lap of time between the harvest of the first winter cereal and the sowing of the second.

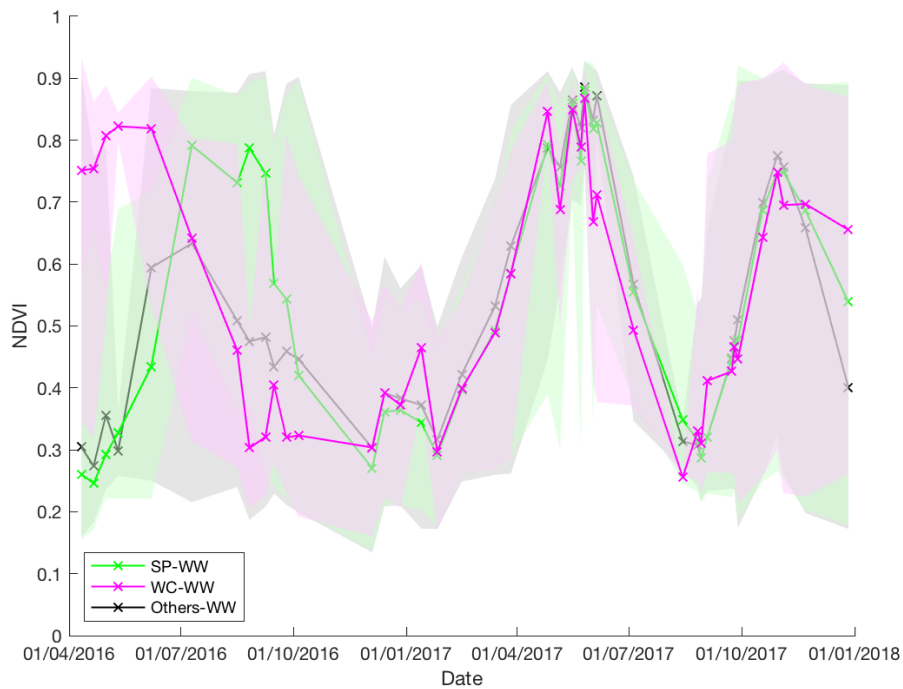


Figure 3.4: Graphical evolution of the NDVI average of winter wheat crops in 2017 preceded by spring crops (green), winter cereals (magenta) or other crops (black) in 2016, showing the 97,5% and 2,5% percentiles around the mean as the coloured patches for each curve

3.1.6 Winter barley

Figure 3.5 shows the evolution of NDVI average of winter barley crops in 2017 and their antecedents in 2016. Winter barley is mostly preceded by winter cereals in magenta (winter wheat and winter barley) and maize in green. Winter barley is minimally preceded by: winter rapeseed, spelt, green beans, fiber flax and peas; regrouped in the "others" category in black. Winter barley is generally planted in October and it is harvested in July, as seen in the right part of Figure 3.5.

Winter barley is mostly preceded by winter wheat in red (Figure 3.5) and winter barley in blue (Figure 3.5). Winter wheat (red) is harvested in September, so there is no time to implement a winter cover crop because winter barley is planted right after in October. Winter barley (blue) is harvested in July. As the following winter barley is planted in October, leaves the possibility for the farmer to implement a winter cover crop between these two winter cereals.

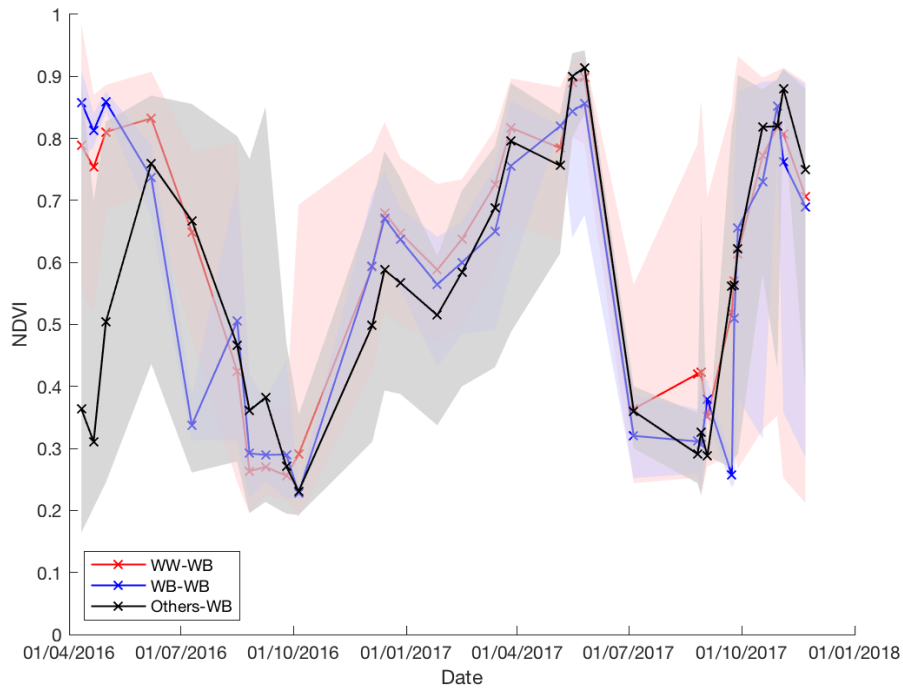


Figure 3.5: Graphical evolution of the NDVI average of winter barley crops in 2017 preceded by winter wheat crops (red), winter barley crops (blue), or other crops (black) in 2016, showing the 97,5% and 2,5% percentiles around the mean as the coloured patches for each curve

3.2 Rule-based Classification

In this section, the results obtained thanks to the rule-based classification method will be presented. First, Figure 3.6 represents the NDVI evolution of each cultural class defined in Section 2.8, with the standard deviation associated to each date. The basic statistics previously presented are derived from the common LPIS map processed in Section 2.5. Nevertheless, the data set first used in this section is the "3 classes map" processed in Section 2.8. It contains 1390 parcels and has a surface of 4344,18 hectares, which accounts for 61% of the LPIS common map (whose surface equals to 7121,57 hectares).

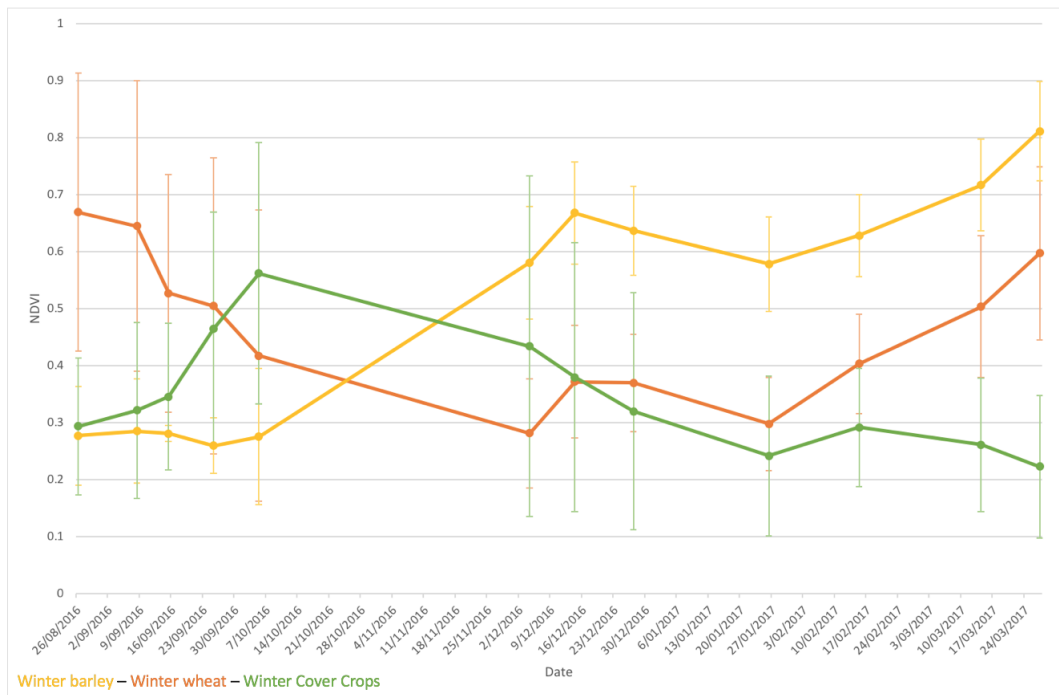


Figure 3.6: Evolution of the mean NDVI averages of winter wheat, winter barley and winter cover crops with their respective standard deviation from August 2016 to March 2017

Figure 3.6 shows that the moments when the standard deviations of each class overlap the least are February and March. October may also be a period when the three classes are well distinguished. Unfortunately, in 2016, the satellite images have a cloud coverage of 97% which was unusable in this case. The temporal windows require a low cloud coverage in order to perform the rule-based classification. Consequently, three dates were identified as potential temporal windows:

- 04/12/2016
- 15/02/2017

– 14/03/2017

The rule-based classification uses decision rules to define NDVI thresholds for different crop types. These decision rules are based on the following pattern:

1. September

- High LAI/NDVI values
 - Spring crop
 - Cover crop (sown in August)
- Low LAI/NDVI values
 - Bare soil
 - Winter cereal (sown in September)
 - Cover crop (sown in September)

2. November

- High LAI/NDVI values
 - Sugarbeet
 - Cover crop (not yet destroyed)
- Low LAI/NDVI values
 - Bare soil
 - Winter cereal (planted in late October or at the beginning of November)

3. March

- High LAI/NDVI values
 - Winter cereal
- Low LAI/NDVI values
 - Bare soil
 - Soil before the sowing of a spring crop

For each date, NDVI thresholds were determined for each class by analyzing the extent of the corresponding standard deviations in Figure 3.6. A first test was carried out using an extra class: the "bare soil". The thresholds for the first test are shown in Table 3.6.

Table 3.6: NDVI thresholds of the four classes for the first test of the rule-based classification method

	04/12/2016	15/02/2017	14/03/2017
Bare soil (0)	0 – 0,19	0 – 0,19	0 – 0,19
Winter wheat (1)	0,19 – 0,38	0,35 – 0,5	0,38 – 0,63
Winter barley (2)	0,5 - max	0,5 - max	0,63 - max
WCC (3)	0,38 – 0,5	0,19 – 0,35	0,19 – 0,38

Then, a second test was realized where the NDVI thresholds (see Table 3.7) were set for three classes (excluding the "bare soil" class).

Table 3.7: NDVI thresholds of the three classes used for the second test of the rule-based classification method

	04/12/2016	15/02/2017	14/03/2017
Winter wheat (1)	0 – 0,38	0,35 – 0,5	0,38 – 0,63
Winter barley (2)	0,5 - max	0,5 - max	0,63 - max
WCC (3)	0,38 – 0,5	0 – 0,35	0 – 0,38

The results from the first and second tests are displayed in Table 3.8 and Table 3.9, respectively. Each test was carried out for a pixel-based and a parcel-based approach. In this method, the accuracy assessment is performed on the overall data set previously defined for the different classification steps (the 3 classes map processed in Section 2.8 counting 1390 parcels). All confusion matrices regarding the tests performed using the rule-based classification method can be found in Appendix C. In Tables 3.8 and 3.9, the results displayed are the overall accuracy (OA) and kappa coefficient. These two metrics come from computed confusion matrices as explained in Section 2.8. The following tables also display the user accuracy and producer accuracy of the WCC (winter cover crops) class.

First of all, what stands out in Tables 3.8 and 3.9 is that the lowest OA and kappa values are obtained for the 4th of December 2016 in both tests and approaches. In fact, at that time, a lot of cropping scenarios are possible. Winter wheat and winter cover crops may have the same spectral signatures if the winter cover crop has already been destroyed. If the winter cover crop is still in place at the beginning of December, the contrast between winter wheat, winter barley and winter cover crops is most likely to be significant. The wide range of possible scenarios is endorsed by the huge size of WCC standard deviation in Figure 3.6. Therefore, distinguishing the three classes at that time of year results quite difficult and as a result the kappa is rather low. The tests for the 15th of February 2017 and the 14th of March 2017 result in a more accurate classification than the 4th of December (Figure 3.6).

Table 3.8: Results from the first test of the rule-based classification method using thresholds defined in Table 3.6

	Pixel-based			Parcel-based		
	4/12	15/02	14/03	4/12	15/02	14/03
OA (%)	42,50	68,39	73,86	42,56	70,77	76,00
kappa (/)	0,173	0,486	0,574	0,176	0,518	0,604
UA _{WCC} (%)	5,33	63,38	73,37	5,43	64,36	75,06
PA _{WCC} (%)	1,50	82,71	91,55	1,63	82,89	91,64

Table 3.9: Results from the second test of the rule-based classification method using thresholds defined in Table 3.7

	Pixel-based			Parcel-based		
	4/12	15/02	14/03	4/12	15/02	14/03
OA (%)	56,25	68,50	73,86	56,06	70,77	76,00
kappa (/)	0,211	0,487	0,574	0,208	0,518	0,604
UA _{WCC} (%)	5,32	62,45	73,36	5,43	64,36	75,06
PA _{WCC} (%)	1,50	83,04	91,55	1,63	82,89	91,64

Then, the differences between the first and the second test are not significant. The only notable difference is the increase of accuracy for the 4th of December 2016 from 4 classes to 3 classes. It appears that the accuracy is improved when the "bare soil" class is integrated into the class having the lowest NDVI thresholds. Overall, the accuracy increases a little for the tests with the parcel-based approach.

However, when looking at the user accuracy and producer accuracy of the WCC class in Tables 3.6 and 3.7, they are similar for the first and the second tests with the pixel-based approach and the parcel-based approach, respectively. Thus, isolating the bare soils into one class do improve the classification of the WCC samples. Another observation that stands out is that UA and PA are very low in each test on the 4th of December 2016. Indeed, distinguishing winter cover crops from winter cereals is difficult at that time. Winter cover crops can either be destroyed or still place. Whereas winter barley is developing itself (sown in October) and winter wheat has just been sown in November. Therefore, using thresholding to discriminate winter cover crops from winter cereals in early December is not appropriate.

As the test with 3 classes (winter wheat, winter barley and winter cover crops) with a parcel-based approach was the one that provided the best results, it was also carried out for other dates when NDVI thresholds could be easily determined:

- 27th of December 2016
- 26th of January 2017
- 27th of March 2017

Table 3.10 presents the thresholds determined for the additional dates.

Table 3.10: NDVI thresholds of the three classes for the additional dates

	27/12/2016	26/01/2017	27/03/2017
Winter wheat (1)	0,32 – 0,5	0,27 – 0,4	0,35 – 0,72
Winter barley (2)	0,5 – max	0,4 – max	0,72 – max
WCC (3)	0 – 0,32	0 – 0,34	0 – 0,35

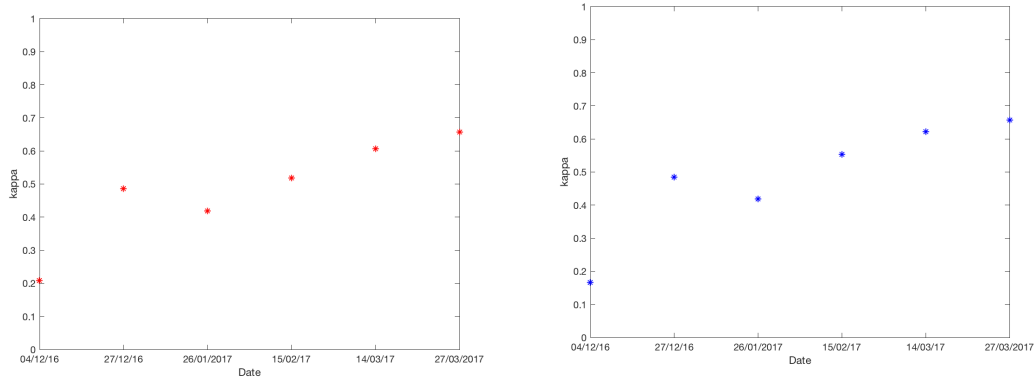
Then, Table 3.11 compares the results of the test with three classes or with the 3 classes extended data set (see details below), and with a parcel-based approach for all the dates previously identified. These results are also presented in Figure 3.7.

Table 3.11: Results of the parcel-based test for all the dates previously identified for the three classes and the three classes using the extended parcel data set

		4/12	27/12	26/01	15/02	14/03	27/03
3 classes	OA (%)	56,06	68,33	62,35	70,77	76,00	78,85
	kappa (/)	0,208	0,486	0,418	0,518	0,604	0,657
	UA _{WCC} (%)	5,43	57,46	58,69	64,36	75,06	90,07
	PA _{WCC} (%)	1,63	69,81	78,65	82,89	91,64	93,19
3 classes (extended)	OA (%)	49,55	68,28	62,43	73,22	77,85	78,85
	kappa (/)	0,166	0,485	0,419	0,553	0,632	0,657
	UA _{WCC} (%)	8,85	57,46	58,69	72,15	80,97	90,07
	PA _{WCC} (%)	1,97	69,81	78,65	85,24	91,52	93,19

Previously, the additional dates were not included in the rule-based classification because the standard deviations of the three classes of interests overlapped too much (Figure 3.6), which means that these classes were less easy to discriminate at the additional dates. However, in order to assess the global accuracy of the rule-based classification method, a comparison of this method at all dates was performed. The results of this test are gathered in Table 3.11 in the "3 classes" row. In this case, OA, kappa coefficients, UA and PA of the WCC class tend to increase with time. Figure 3.7a also corroborates this observation. In fact, discriminating winter cover crops from winter cereals becomes more and more easy when approaching March. However, some agricultural plots are already sown with sugarbeet on the 14th of March, so it is rather late for purposes of an early

warning system (see Section 3.5). The metrics providing the poorer results are the ones of the 4th of December. This is consistent with the fact that it is difficult to distinguish winter cover crops from winter cereals is difficult at that time because winter cover crops can either be destroyed or still place. Whereas winter barley is developing itself (sown in October) and winter wheat has just been sown in November. So, as previously mentioned, early December might not be an appropriate period to discriminate winter cover crops from inter cereals.



(a) Evolution of the κ obtained for the rule-based classification with three classes and a parcel-based approach (b) Evolution of the κ obtained for the rule-based classification with three classes extended parcel data set and a parcel-based approach

Figure 3.7: Graphs representing the kappa evolution of the obtained results in the last rule-based classification tests

In order to assess the differences between winter cover crops and the other types of crops that precede spring crops, it is necessary to distinguish the NDVI time series of both kinds of crops. In Figure 3.8, the green line represents the NDVI average of winter cereals (winter wheat and winter barley) in 2016 followed by sugarbeet, maize or potato in 2017 (same green curve as in Figure 3.6). On the other hand, the blue curve stands for the other crops that preceded sugarbeet, maize and potato in 2016. These other crops include linen, maize, sugarbeet, potato, rapeseed, chicory, pea, etc, i.e. all other crops besides winter wheat and winter barley. The difference between these two curves is that in the case of the green curve, it is most likely that there will be a winter cover crop between the winter cereal and the spring crop (see Section 1.2 in Chapter "State of Art"). By contrast, in the blue curve, it is less likely to have a winter cover crop between the crops regrouped in the "other crops" category and spring crops. This is corroborated by this chart where the blue curve drops when the green curve increases at the time of the growth of winter cover crops. However, the lack of usable images in October-November does not allow to fully verify that statement.

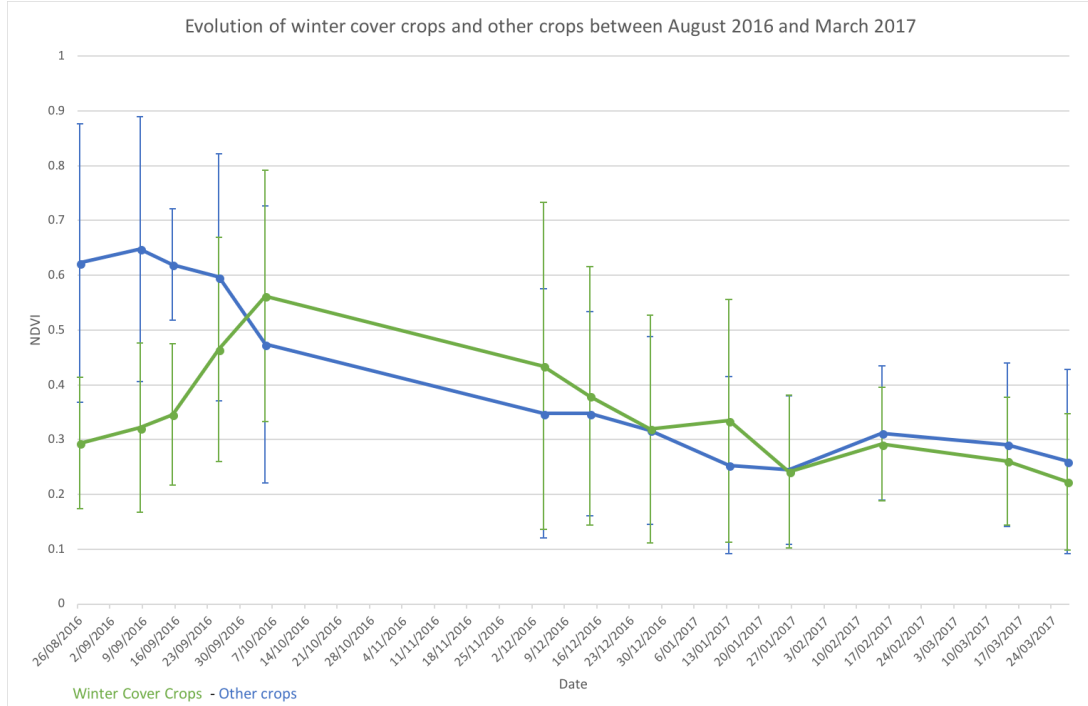


Figure 3.8: Evolution of the NDVI averages of winter cover crops and other crops each followed by maize, sugarbeet or potato, with their respective standard deviation from August 2016 to March 2017

The final test using rule-based classification consists in a parcel-based approach using all of the previously identified dates, three classes and a larger data set. In the previous tests the rule-based classification was carried out using three classes: winter wheat, winter barley and spring crops (sugarbeet, maize, potato) preceded by winter cereals. In this last test, the other antecedents of maize, sugarbeet and potato were added to the third class (the "WCC" class) in order to assess the impact of these antecedents while using the same decision rules (NDVI thresholds) as before. This "3 classes extended map" contains 1652 parcels and has a surface of 4944,26 hectares, which accounts for 69,4% of the surface of the common LPIS map (whose surface equals to 7121,57 hectares). In this case, the accuracy results are also presented in Table 3.11 in the "3 classes extended" row. As for the previous test, metrics seem to increase with time (Figure 3.7b). Except for the December dates, metrics tend to be a little more higher than with the 3 classes previously used. Extending the data set by adding the other antecedents of the spring crops (maize, sugarbeet, potato) had a positive impact on the classification, especially in February and March when the distinction between winter cover crops, i.e. spring crops to be, and winter cereals is more achievable. However, the use of the dates in March can, once again, be questioned due to the growth cycles and plantation of some spring crops (e.g. sugarbeet). This topic will be further discussed in Section 3.4.

3.3 Maximum Likelihood Classification

In this section, the different Maximum Likelihood classifications carried out will be explained and their results will be presented. All the following tests were performed using a pixel-based approach at parcel scale, i.e. every pixel value was taken into account by the ML classifier. The sample data set used in this section is the "3 classes map" processed in 2.9. It contains 1390 parcels and has a surface of 4344,18 hectares, which accounts for 61% of the surface of the common LPIS map (whose surface equals to 7121,57 hectares). The accuracy assessment in each test was performed using the validation data set isolated as described in Section 2.9. The accuracy of the calibration was not assessed in this part of the master thesis.

First, different kinds of features were used in the Maximum Likelihood classifications for different dates (previously determined) and in different combinations. Table 3.12 shows the different combinations of features carried out in the classification tests. For each combination, a number was assigned in order to make the display of the results easier.

Table 3.12: Types of tests in the Maximum Likelihood Classification

Number	Type of feature
1	NDVI
2	10 bands
3	NDVI + 10 bands
4	NDVI + 10 bands + SAVI
5	NDVI + 10 bands + NDWI
6	NDVI + 10 bands + Brightness
7	NDVI + 10 bands + SAVI + NDWI
8	NDVI + 10 bands + SAVI + NDWI + Brightness

The first part of the Maximum Likelihood classifications is presented in Table 3.13. The first column of Table 3.13 stands for the number of classes to classify according to the maximum likelihood. They were determined in Section 2.9.1: winter wheat, winter barley and winter cover crops (WCC). In the second column, the number of the combined features (Table 3.12) is displayed. Then, each column stands for a date and a cross marks which date has been used for each classification test. Finally, the last two columns shows the overall accuracy (OA) and the kappa of the classification test.

Table 3.13: Results from the Maximum Likelihood classification method (0); UA and PA are of the WCC class (25/09/2016, 04/12/2016, 27/12/2016, 15/02/2017, 14/03/2017)

Classes	Nbr	25/09	04/12	27/12	15/02	14/03	OA (%)	κ (/)	UA (%)	PA (%)
3	1	X					41,57	0,172	43,98	47,06
3	1		X				76,04	0,519	97,38	39,22
3	1			X			78,84	0,618	85,58	55,39
3	1				X		69,93	0,496	59,01	79,37
3	1					X	73,82	0,567	72,49	89,67
3	2	X					61,90	0,423	78,49	61,96
3	2		X				79,98	0,589	92,57	48,71
3	2			X			78,91	0,567	91,18	45,63
3	2				X		85,78	0,724	87,53	73,77
3	2					X	92,21	0,854	89,44	93,46
3	3	X					61,66	0,422	78,59	61,30
3	3		X				80,39	0,597	93,64	48,76
3	3			X			85,58	0,717	94,45	67,34
3	3				X		91,04	0,831	91,24	89,91
3	3					X	91,41	0,84	89,88	93,22
3	3		X	X			81,61	0,623	96,79	51,75
3	3		X	X	X		82,39	0,639	96,79	53,53
3	3			X	X	X	94,92	0,903	98,37	91,88
3	3		X		X	X	94,27	0,89	97,78	90,34
3	3		X	X	X	X	91,8	0,839	98,36	81,08

The first classifications tests were carried out using single-date imagery, i.e. only one date and one feature at a time. In this case, the 25th of September 2016 images were used instead of the 5th of October images because the latter contained too many clouds (90% cloud coverage) which was unusable for the Maximum Likelihood classifier. When using the NDVI images for the classification, the September image is the least accurate with a kappa coefficient of 0,172, which is quite low. In fact, at this date, distinguishing winter cereals from winter cover crops results rather difficult for several reasons. First, winter barley might be already planted but not developed yet, so it may be mistaken for a bare soil. Second, winter wheat is not planted yet. So, it is most likely that these parcels are also mistaken for bare soils. And finally, some winter cover crops may already be developed and others not. This results in a huge standard deviation (Figure 3.6) leading to a low kappa coefficient and a low OA.

The OA and kappa coefficients are higher when using the 10 bands images than with the NDVI images. The reason why the metrics are better with the 10 band images is the number of bands. This feature provides a large number of bands, so it covers a larger part of the electromagnetic spectrum, whereas the NDVI images have only one band based on two other bands (NIR and red). The information is therefore more limited in the NDVI

images.

The last classification tests using single-date imagery was performed using a combination of two features: NDVI and 10 bands images. Combining these images increased the overall accuracy and kappa coefficient of three dates: 04/12/2016, 27/12/2016 and 15/02/2016. But the metrics remained the same or close to the values obtained using the 10 band image classification for the other two dates: 25/09/2016 and 14/03/2017. So, the addition of the NDVI images contributed to increase the information used by the Maximum Likelihood classifier, hence raising the OA and kappa coefficient.

Figure 3.9 presents the graphical evolution of the accuracy assessment results mentioned above. It clearly shows that the combination of NDVI images and 10 band images provides more accuracy than the separate images when using single-date imagery.

The lower part of Table 3.13 provides the accuracy assessment results of Maximum Likelihood classification achieved using multi-date imagery and the combination of NDVI images and 10 band images. In these tests, the 25th of September 2016 image was left out because it was not adding significant information to improve the accuracy of the Maximum Likelihood classification. Then, different combinations of dates were tested, still using both NDVI and 10 band images, in order to assess the best combination of dates. As shown in Table 3.13, the Maximum Likelihood classification test with the highest accuracy is the one combining 27/12/2016, 15/02/2017 and 14/03/2017, with an overall accuracy of 94,92 and a kappa of 0,903.

Table 3.13 also shows the user accuracy and the producer accuracy of the WCC class, i.e. the crop type class that needs to be most correctly classified. In the tests using one feature (Nbr 1 & 2 in Table 3.13), both accuracies tend to increase when getting closer to March. Nevertheless, on the 15th of February 2017, UA and PA tend to be a little lower than for the previous and the following dates. And, for this date in test Nbr 1 (Table 3.13), UA is lower than PA which means that the samples were less correctly classified among all samples than with respect to the validation map. In the tests using single-date imagery (Nbr 1, 2 & 3 in Table 3.13), PA of the WCC class is lower than UA for the first three dates. So, for these three dates, the samples were less correctly classified with respect to the validation map. In the lowest part of Table 3.13 regarding multi-date imagery, UA and PA are the highest (98,37% and 91,88% respectively) when the OA and κ coefficient are highest too. What stands out in this part of the table is that for the last multi-date imagery test (04/12/2016 + 27/12/2016 + 15/02/2017 + 14/03/2017), the PA is lower than the PA's of the two previous multi-date imagery tests in the table. So,

using both December dates results in more incorrectly classified samples with respect to the validation map.

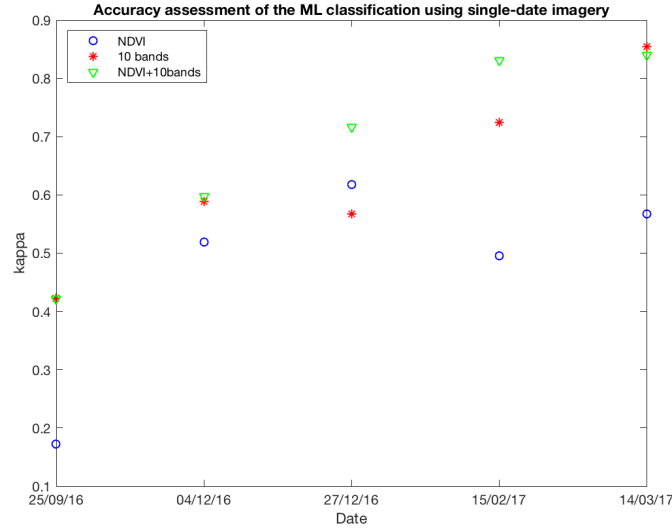


Figure 3.9: Results of the accuracy assessment of the Maximum Likelihood classification using single-date imagery and two features: NDVI images and 10 bands images

The third class, winter cover crops, is composed of the spring crops (maize, sugarbeet and potato) preceded by either winter wheat or barley. Thus, this class regroups different possible behaviours which, therefore, increases its variance. In order to reduce this variance, to increase the accuracy of the classification at issue, the third class was split into two sub-classes according to the winter cereal preceding the spring crop, resulting into four classes to perform the Maximum Likelihood classification. Table 3.14 displays these accuracy assessment results. UA_1 and PA_1 correspond the user accuracy and producer accuracy of the WCC sub-class where maize, sugarbeet and potato are preceded by winter wheat (WCCa). Whereas, UA_2 and PA_2 correspond the user accuracy and producer accuracy of the WCC sub-class where maize, sugarbeet and potato are preceded by winter barley (WCCb).

What stands out in the table is that the split of the "winter cover crops" class into two sub-classes did not increase the overall accuracy and kappa as predicted. Also, even though the UA_1 is rather high in every case, the PA_1 accuracy is lower than the UA_1 in each case. The Maximum Likelihood classifier finds it hard to correctly classify the spring crops preceded by winter wheat samples with respect to the validation map. UA_2 and PA_2 are both rather low. So, the ML classifier has difficulties to classify the samples when spring crops are preceded by winter barley. The maximum likelihood classifier finds it

easier to classify the parcels in the study zone according to three classes rather than four.

Table 3.14: Results from the Maximum Likelihood classification method (1); UA and PA are of the WCC sub-classes (04/12/16 & 27/12/16 & 15/02/17 & 14/03/17)

Classes	Nbr	04/12	27/12	15/02	14/03	OA (%)	κ (/)	UA ₁ (%)	PA ₁ (%)	UA ₂ (%)	PA ₂ (%)
4	3	X				76,94	0,547	80,17	42,42	26,00	23,74
4	3		X			73,82	0,550	82,83	26,98	13,75	59,51
4	3			X		75,24	0,581	80,03	34,36	15,88	59,14
4	3				X	73,6	0,559	68,26	29,79	14,95	69,01
4	3	X	X			78,15	0,575	84,97	42,38	27,04	32,97
4	3	X	X	X		78,41	0,585	85,32	41,84	22,24	33,69
4	3		X	X	X	78,23	0,628	83,67	32,84	15,28	64,84
4	3	X		X	X	80,94	0,668	84,85	44,81	16,75	56,98
4	3	X	X	X	X	80,06	0,652	85,82	39,71	16,45	58,74

Finally, a last series of Maximum Likelihood classification tests was performed to evaluate the impact of several coupled features. As mutli-date imagery resulted in producing better classifications than single-date imagery, these tests were set up using four dates: 04/12/2016, 27/12/2016, 15/02/2017 and 14/03/2017. Also, this series of tests were performed for the association of these four dates providing the best results (Table 3.15):

- 27/12/2016 + 15/02/2017 + 14/03/2017 (a)
- 04/12/2016 + 15/02/2017 + 14/03/2017 (b)
- 04/12/2016 + 27/12/2016 + 15/02/2017 + 14/03/2017 (c)

Moreover, all tests were carried out using NDVI and 10 bands images combined with other indexes. These different combinations are referenced by the number column in Table 3.15, and each combination is described in Table 3.12.

In combination number 4, the Soil Adjusted Vegetation Index is added to the other two features. This index is designed to correct the effect of soil brightness when there is a low vegetation cover. As it is often the case in autumn and winter, the addition of the SAVI improved the accuracy assessment results (Table 3.15). Then, in combination number 5 and number 6, the same tests were carried out using the Normalized Difference Water Index and the Brightness Index, respectively, in addition to the NDVI and 10 bands images. What stands out in Table 3.15 is that the OA and kappa in combination 5 and 6 are a bit lower than when using the Soil Adjusted Vegetation Index. In the last two tests (7 and 8), the Maximum Likelihood Classification was carried out using four and five features. In every test, what stands out in this table (as in Table 3.13) is that the PA of the last combination of dates (c) is lower than for the previous combinations of dates (a & b). Another observation that can be made by observing the table is that UA's are the highest when using Brightness (test Nbr 6, Table 3.15) as a third feature. But

PA's are higher when using SAVI as a third feature, also providing high UA's. It appears that using more than three features does not improve the accuracy assessment results. In fact, the best accuracy (OA, κ , UA, PA) results are obtained combining NDVI, 10 bands and SAVI images. Using SAVI as a third feature provides the most correctly classified samples among all samples and with respect to the validation map of the WCC class.

Table 3.15: Results from the Maximum Likelihood classification method (2); UA and PA are of the WCC class

Classes	Nbr	04/12/16	27/12/16	15/02/17	14/03/17	OA (%)	κ (/)	UA (%)	PA (%)
3	4		X	X	X	96,07	0,925	98,00	95,96
3	4	X		X	X	95,59	0,916	97,85	94,98
3	4	X	X	X	X	94,84	0,901	98,14	90,26
3	5		X	X	X	95,05	0,905	97,62	92,27
3	5	X		X	X	94,77	0,900	97,43	91,79
3	5	X	X	X	X	92,99	0,863	97,80	84,81
3	6		X	X	X	95,57	0,915	98,45	93,69
3	6	X		X	X	95,2	0,908	98,18	93,20
3	6	X	X	X	X	92,65	0,863	98,49	84,42
3	7		X	X	X	95,75	0,919	97,28	95,20
3	7	X		X	X	95,41	0,913	97,39	94,62
3	7	X	X	X	X	94,77	0,899	97,54	91,05
3	8		X	X	X	96,02	0,924	97,23	95,44
3	8	X		X	X	95,58	0,916	97,44	94,93
3	8	X	X	X	X	95,07	0,905	97,63	91,52

3.4 General discussion

First, a comparison between rule-based classification and maximum likelihood classification using the same patterns was made in Table 3.16: using NDVI images, three classes and a pixel-based approach. What stands out in Table 3.16 is that the maximum likelihood classifier is much more efficient than the rule-based method regarding the 4th of December 2016. Using training areas instead of an automatic method is more efficient at this time of year. With regard to the other two dates, the kappa are very similar.

Table 3.16: Comparison of the kappa obtained using three classes and a pixel-based approach in rule-based and maximum likelihood classification

	04/12/2016	15/02/2017	14/03/2017
Rule-based (kappa [/])	0,221	0,487	0,574
Maximum Likelihood (kappa [/])	0,519	0,496	0,567

Then, the rule-based classification method provided the best accuracy results with three classes and a parcel-based approach where each parcel was given a value based on the average of the pixels inside the corresponding parcel. In this case, the kappa coefficients reached 0,486; 0,518 and 0,604 respectively for 27/12/2016, 15/02/2017 and 14/03/2017 (Table 3.11).

The rule-based classification method was developed to assess the accuracy of an automatic classification method based on decision rules, i.e. NDVI thresholds defined depending on the land cover classes. In the best cases, this method reaches an overall accuracy of 70 to 76 %, which means about 30% of errors. One of the strengths of the rule-based methodology is that as time passes by, the accuracy increases. However, that can also be a weakness because one of the objectives of the early warning system would be to identify the hazardous areas the earliest possible, preferably before March. But in this case, accuracy increases when coming closer to March, so the rule-based classification method is less accurate before March and involves more classification errors. Another difficulty with this method is the replicability over the years due to inter-annual variability of the different types of crops. As each year is different, the NDVI thresholds may vary from one year to another making the replicability rather difficult, perhaps impossible.

The Maximum Likelihood classification method was the effective when combining three features (NDVI, 10 bands and SAVI images) and different dates (in this case: 27/12/2016 + 15/02/2017 + 14/03/2017) suitable to distinguish the study zone parcels according to the three classes defined (winter wheat, winter barley and winter cover crops) and using a pixel-based approach. It happens to reach an accuracy of 96% and a high kappa coef-

ficient of 0,925.

Although the outcome is satisfying, there are some limitations to both methods developed in this master thesis. In the rule-based method, single-date imagery was used. Even though Yang *et al.* (2011) promoted the use of single-date imagery in crop identification method, this master thesis proves that multi-date imagery provides better accuracy results. In fact, single-date imagery can be used and can provide high accuracy results in order to identify crops depending if the optimum crop discrimination period is well chosen (Yang *et al.*, 2011). So, the use of the Maximum Likelihood method has demonstrated that multi-date imagery is more appropriate in crop identification and classification. In fact, using multi-date imagery, i.e time series, in order to discriminate crop types offers a wide range of opportunities such as a larger data set, the possibility of combining key plant growth periods in the classification and a better characterization of land cover types (Gomez *et al.*, 2016).

One of the purposes of the ML classification was to evaluate the difference between an automatic method (rule-based) and one using training areas and probabilities (ML). As demonstrated by the best results in each method, the Maximum Likelihood classification provides the more error-free performance, reaching 96% of overall accuracy in the best case. It is true that the Maximum Likelihood classification is simple to apply and easy to understand (Gomez *et al.*, 2016 ; Defourny, 2017a). Nevertheless, this methodology also comes with disadvantages. It is a parametric algorithm, it assumes a normal distribution of the data and it requires a large amount of samples (Gomez *et al.*, 2016 ; Defourny, 2017a).

The rule-based and ML classification methods are rather highly sensitive to the lack of data due to cloud coverage. Developing a classification method using the slope of NDVI values over a period of time instead of the NDVI values of every single date could provide the answer the replicability difficulties encountered in the rule-based and ML classifications, but could also cope with the lack of data when necessary. As a matter of fact, NDVI values on any given date may vary from one year to another but the general trend (i.e. slope) over a period of time may be less variable. This could be achieved by using the slope of NDVI values calculated by linear regression over a given period of time. Another alternative to reduce the sensitivity to the lack of data would be to apply object-based image analysis (OBIA) to extract crop types from multi-series data (Belgiu *et al.*, 2018). OBIA is performed prior to the classification (supervised or unsupervised) and is an iterative method that uses satellite imagery segmentation to produce image objects then assigned the defined crop classes using supervised or unsupervised classifi-

cation (Blaschke, 2010). The use of the Random Forest algorithm might also provide an opportunity to diminish the sensitivity due to the lack of data available (Gomez *et al.*). It provides high classification accuracy and has proven its efficiency in various studies (Belgiu *et al.*, 2018; Hao *et al.*, 2015; Pelletier *et al.*, 2016)

In every method developed, dates in March have been used. However, on the 14th of March, a lot of plots are already sown with sugarbeet. So, this date is rather late for an early warning system. It is important to note that in early 2017, Sentinel-2A was the only satellite of the constellation in orbit as Sentinel-2B was launched in March 2017. In the following years, there will be more images available in each month thanks to the two satellites and this will provide more information at the key periods: October-December and January-February. Also, one of the biggest difficulties in this master thesis was that there were no images available at the period when the differences between winter cereals and winter cover crops are more numerous (October-December) due to high cloud coverage and the presence of only one satellite in orbit of the Sentinel-2 constellation.

Conclusion & Perspectives

With the prospect of discriminating winter cereals from winter cover crops during winter season, different classification methods were elaborated in this master thesis: a rule-based classification method and a method based on the use of the Maximum Likelihood classifier.

The first method developed in this master thesis was based on decision rules defining NDVI thresholds according to the land cover class (winter wheat, winter barley and winter cover crop): the rule-based classification. This automatic method was most efficient using the three classes just mentioned and a parcel-based approach, in which every parcel in the study zone is given a single NDVI value by operating the pixels average. Then, the Maximum Likelihood classification was proceeded using single-date and multi-date imagery, single features and combinations of features: NDVI, 10 bands, SAVI, NDWI and Brightness images. The Maximum Likelihood classification gave better accuracy results than the rule-based classification with an overall accuracy reaching 96,07% and a kappa coefficient of 0,925 in the case of:

- a combination of three features: NDVI, 10 band and SAVI images;
- a combination of three dates: 27/12/2016, 15/02/2016 and 14/03/2016.

This test also provided high user accuracy and producer accuracy of the WCC class which was the one that needed to be the most correctly classified. Then, a last methodology was developed in order to ensure a lower sensitivity to the lack of data and a better robustness of the method. This methodology used the NDVI slopes as input images in the Maximum Likelihood classification. In order to expand the use of this method, it would be necessary to perform additional tests using the slopes of the 10 band and SAVI images values in order to optimize the accuracy of the method results.

Even though the achieved results meet the main objective set in this master thesis, it is necessary to put theses achievements into perspective in order to get a step closer to creating an early warning system useful to the municipalities with muddy flood problems. It would be interesting to apply the same methods for the winter season in 2017-2018 and

2018-2019 in order to assess the replicability of these methods. As the Maximum Likelihood classifier is less widely used nowadays, it might be worthwhile to use a different classification method. For example, the Random Forest algorithm provides stability and efficiency to obtain a high degree of accuracy. It also handles large amounts of data and is more robust than the ML classifier. This algorithm is the one used to process the products provided by the Sentinel-2 for Agriculture program. In this master thesis, the reference map used in the different classifications was processed using the LPIS maps from 2016 and 2017. However, as previously mentioned, this map is only available for year n in January of year $n+1$. An alternative could be to use the crop type product provided by Sentinel-2 for Agriculture as a reference map. Then, a significant difficulty in this master thesis was the lack of images during October-December 2016. Therefore, it may be useful to perform the same tests (rule-based and ML classification) using feature images acquired during October, November and December 2017 (or 2018) in order to evaluate the contribution of these periods to the distinction between winter cereals and winter cover crops.

Finally, The biggest remaining challenge is that the information about the next crop on a given agricultural plot in real time will never be available. It will need some kind of predictions which will involve more errors and more uncertainty.

Bibliography

- [1] Belgiu, M., Csillik, O. (2018). Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sensing of Environment*, 204, 509-523.
- [2] Bertin, P. (2017). *Phytotechnie*. Cours, Université catholique de Louvain, Louvain-la-Neuve.
- [3] Bielders, C., Colard, F., Cordonnier, H., Degré, A., Destain, M.-F., Feltz, N., Maugnard, A. (2011a). *Convention de recherche d'intérêt général Gestion Intégrée Sol Erosion Ruissellement : Annexes*. Convention GISER, Service Pubic de Wallonie, Direction générale opérationnelle agriculture, ressources naturelles et environnement, Département de la ruralité et des cours d'eau, Direction du développement rural.
- [4] Bielders, C., Colard, F., Cordonnier, H., Degré, A., Destain, M.-F., Feltz, N., Maugnard, A. (2011b). *Convention de recherche d'intérêt général Gestion Intégrée Sol Erosion Ruissellement : Rapport Final*. Convention GISER, Service Pubic de Wallonie, Direction générale opérationnelle agriculture, ressources naturelles et environnement, Département de la ruralité et des cours d'eau, Direction du développement rural.
- [5] Bielders, C., Ramelot, C., Persoons, E. 2003. Farmer perception of runoff and erosion and extent of flooding in the silt-loam belt of the Belgian Walloon Region. *Environmental Science & Policy*, 6 (1), 85-93.
- [6] Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65 (1), 2-16.
- [7] Boardman, J., Vandaele, K. (2010). Soil erosion, muddy floods and the need for institutional memory: Soil erosion, muddy floods and institutional memory. *Area*, 42 (4), 502-513.
- [8] Creative Commons Attribution-ShareAlike 3.0 license. *QGIS - The Leading Open Source Desktop GIS*. En ligne <https://www.qgis.org/en/site/about/index.html>, consulted on the 26th of March 2018.

- [9] Defourny, P. (2017). *Géomatique appliquée : Télédétection aéroportée et spatiale*. Cours, Université catholique de Louvain, Louvain-la-Neuve.
- [10] Defourny, P. (2017a). Land cover mapping and monitoring. In: J. Decliné (ed.), *Handbook on Remote Sensing for Agricultural Statistics* (Chapter 2). Handbook of the Global Strategy to improve Agricultural and Rural Statistics (GSARS): Rome.
- [11] Defourny, P., Radoux, J. (2007). A quantitative assessment of boundaries in automated forest stand delineation using very high resolution imagery. *Remote Sensing of Environment*, 110 (4), 468-475.
- [12] European Climate Adaptation Platform. (2015). *Establishment of early warning systems*. En ligne <http://climate-adapt.eea.europa.eu/metadata/adaptation-options/establishment-of-early-warning-systems>, consulted on the 17th of April 2018.
- [13] European Spatial Agency (ESA1). *Copernicus: observing the earth*. En ligne https://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Overview4, consulted on the 9th of April 2018.
- [14] European Spatial Agency (ESA2). *Introducing Sentinel-2*. En ligne https://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Sentinel-2/Introducing_Sentinel-2, consulted on the 9th of April 2018.
- [15] European Spatial Agency (ESA3). *Sentinel Online: Sentinel-2*. En ligne <https://sentinel.esa.int/web/sentinel/missions/sentinel-2>, consulted on the 20th of November 2017.
- [16] Evrard, O., Biélders, C., Vandaele, K., van Wesemael, B. (2007a). Spatial and temporal variation of muddy floods in central Belgium, off-site impacts and potential control measures. *CATENA*, 70 (3), 443-454.
- [17] Evrard, O., Heitz, C., Liégeois, M., Boardman, J., Vandaele, K., Auze, A.-V., van Wesemael, B. 2010. A comparison of management approaches to control muddy floods in central Belgium, northern France and southern England. *Land Degradation & Development*, 21 (4), 322-335.
- [18] Evrard, O., Persoons, E., Vandaele, K., van Wesemael, B. (2007b). Effectiveness of erosion mitigation measures to prevent muddy floods: A case study in the Belgian loam belt. *Agriculture, Ecosystems & Environment*, 118, 149-158.
- [19] Gao, B.-C. (1996). NDWI - A Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water from Space. *Remote Sensing of Environment*, 58 (3), 257-266.

- [20] Gomez, C., White, J.C., Wulder, M.A. (2016). Optical remotely sensed time series data for land cover classification: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 116, 57-72.
- [21] Gourre, J.-P. (2011). *L'optique dans les instruments : Généralités*. En ligne <https://books.google.be/books?id=zn2TY-PkaPsC&pg=PA203&lpg=PA203&dq=classification+seuillage&source=bl&ots=RRkkKFAj2p&sig=JgUSOAhtgBQEXPhkBRs4VVKfF48&hl=nl&sa=X&ved=0ahUKEwjxuv7nr77aAhXLyaQKHcf0ByAQ6AEIbDAF#v=onepage&q=classification%20seuillage&f=false>
- [22] Hao, P., Zhan, Y., Wang, L., Niu, Z., Shakir, M. (2015). Feature selection of time series MODIS data for early crop classification using random forest: a case study in Kansas, USA. *Remote Sensing*, 7, 5347-5369.
- [23] Hovis, C., Hu, Q., Song, Q., Tang, H., Wu, W., Xiang, M., Zhou, Q. (2017). In-Season Crop Mapping with GF-1/WFV Data by Combining Object-Based Image Analysis and Random Forest. *Remote Sensing*, 9 (11), pp.19.
- [24] Huete, A. R.(1988). A Soil-Adjusted Vegetation Index. *Remote Sensing of Environment*, 25, 295-309.
- [25] Janssens, A-S., Marot, J., Vienne, M.-F. (2012). Dossier érosion et coulées de boue.
- [26] Ledant, M. 2007. *Les indices de végétation*. (Travail de fin d'études de Bachelier, Université de Liège, Liège). Online: <https://fr.slideshare.net/MartinLedant/tfe-les-indices-de-vgtation>
- [27] Leteinturier, B., Tychon, B., Oger, R. (2007). Diagnostic agronomique et agro-environnemental des successions culturales en Wallonie (Belgique). *BASE*, 11 (1), 27-38.
- [28] Liu, C., Frazier, P., Kumar, L. (2007). Comparative assessment of the measures of thematic classification accuracy. *Remote Sensing of Environment*, 107 (4), 606-616.
- [29] MAPIR CAMERA. (2018). *Calibrating Images in Fidji (ImageJ) with MAPIR Plugin*. En ligne <https://www.mapir.camera/pages/calibrating-survey-images>, consulted on the 19th of April 2018.
- [30] NATAGRIWAL. Les Méthodes Agri-environnementales et Climatiques : Tournière enherbée. En ligne <https://www.natagriwal.be/fr/mesures-agro-environnementales/tourniere-enherbee>, consulted on the 21th of May 2018.

- [31] Nitrawal. *CIPAN : Cultures Intermédiaires Pièges A Nitrates*. En ligne, <http://www.agreau.be/uploads/documents/57//fichecipan2015page.pdf>, consulted on the 1st of July 2018.
- [32] Pelletier, C., Valero, S., Inglada, J., Champion, N., Dedieu, G. (2016). Assessing the robustness of random forests to map land cover with high resolution satellite image time series over large areas. *Remote Sensing of Environment*, 187, 156–168.
- [33] Peña-Barragán, José M., Ngugi, Moffat K., Plant, Richard E., Six, J. (2011). Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sensing of Environment*, 115, 1301 - 1316.
- [34] PROTECT'eau. (2018). *Couverts Végétaux*. Online, <https://protecteau.be/fr/nitrate/agriculteurs/couvert>, consulted on the 15th of July 2018.
- [35] Quadratic SPRL. (2016). *Le suivi de l'évolution des cultures en utilisant l'indice NDVI*. En ligne <http://www.quadratic.be/le-suivi-de-levolution-des-cultures-en-utilisant-lindice-ndvi/>, consulted on the 21th of November 2017.
- [36] Rouse, J.W., Jr., Haas, R.H., Schell, J.A., Deering, D.W. (1974). Monitoring vegetation systems in the great plains with ERTS. Presented at Third Earth Resources Technology Satellite-1 Symposium, NASA, Washington, DC, 309–317.
- [37] Sentinel-2 for Agriculture (Sen2Agri). *The Sen2-Agri System*. Online, <http://www.esa-sen2agri.org>, consulted on the 7th of March 2018.
- [38] Schmedtmann, J., Campagnolo, M. (2015). Reliable Crop Identification with Satellite Imagery in the Context of Common Agriculture Policy Subsidy Control. *Remote Sensing*, 7 (7), 9325-9346.
- [39] Urdiain-Arraiza, M. (2018). *Évaluation de deux méthodes de prédiction de l'occupation du sol basées sur les rotations de cultures dans une perspective de prédiction d'inondations boueuses*. (Master thesis). Université catholique de Louvain, Louvain-la-Neuve.
- [40] Verstraeten, G., Poesen, J., Goossens, D., Gillijns, K., Bielders, C., Gabriels, D., Ruysschaert, G., Van Den Eeckhaut, M., Vanwalleggem, T., Govers, G. (2006). Chapter 30 : Belgium. In Boardman, J. & Poesen, J. (Eds), *Soil Erosion in Europe* (pp.385-411). England: John Wiley & Sons.
- [41] Verstraeten, G., Poesen, J. 1999. The nature of small-scale flooding, muddy floods and retention pond sedimentation in central Belgium. *Geomorphology*, 29 (3-4), 275-292.

- [42] Yang, C., Everitt, J., Murden, D. (2011). Evaluating high resolution SPOT 5 satellite imagery for crop identification. *Computers and Electronics in Agriculture*, 75 (2), 347-354.
- [43] Yang, Y., Huang, Q., Wu, W., Luo, J., Gao, L., Dong, W., Wu, T., Hu, X. (2017). Geo-Parcel Based Crop Identification by Integrating High Spatial-Temporal Resolution Imagery from Multi-Source Satellite Data. *Remote Sensing*, 9 (12), pp.20.

Early warning system against muddy floods in central Belgium

Presented by Alice Danzin

Summary Central Belgium is one of the most muddy flood affected regions in Europe. One of the main causes of muddy floods is the presence of spring crops (e.g. maize, sugarbeet, potato) in the catchments. In Belgium, winter cover crops are implemented during the intercrop period preceding a spring crop and following a winter cereal. The Land Parcel Identification System provides the information on crop rotations every year in Belgium. Also, remote sensing acquires significant information on vegetation and plant behaviour that are becoming more and more useful in agricultural management and crop identification. With the increasing frequency of muddy flood events, the damage and its costs to municipalities, as well as the psychological stress on people, there is a need to identify hazardous areas in order to implement mitigation measures. To achieve the design of a such an early warning system against muddy floods, it is essential to discriminate winter cereals from winter cover crops during winter season. In order to meet this objective, different procedures were carried out to classify sample plots from a study zone into different classes (winter wheat, winter barley and winter cover crops): a rule-based classification method and a Maximum Likelihood classification method.

By using different combinations of dates and remotely sensed features, what stands out in this master thesis is that using the Maximum Likelihood classifier with a combination of three features (NDVI, 10 band and SAVI images) and a combination of chronological dates (27/12/2016, 15/02/2017 and 14/03/2017) provide the most accurate results and allow the identification of winter cover crops with a high level of accuracy. With the aim of improving the classification methods, it would be interesting to add feature images taken during October and November because the differences between winter cereals and winter cover crops are expected to be really significant during these two months. Then, using a different classification algorithm, like Random Forest, could be a way to improve the robustness and replicability of the classification pattern. The challenge in the building of an early warning system against muddy floods using crop discrimination is that the information about the next crop on a given agricultural plot in real time will never be available. It will need some kind of predictions which will involve more errors and more uncertainty.

Keywords: crop classification, remote sensing, muddy floods, central Belgium.