

Faculté des bioingénieurs

Monitoring of stormwater basins with Sentinel-2 image time series over the province of Walloon Brabant (Belgium)

Auteur :	Caroline Vassart
Promoteurs :	Professeur Pierre Defourny (UCL/ELIe/ENGE) Professeur Mathieu Javaux (UCL/ELIe/GERU)
Lecteurs :	Docteur Julien Radoux (UCL/ELIe/ENGE) Damien Delforge (UCL/ELIe/GERU)
Année Académique :	2018 - 2019
Bioingénieur :	Sciences et technologies de l'environnement.

Acknowledgments

Through these acknowledgements, I would like to express my heartfelt gratitude to all those who contributed to the achievement of this research project and subsequent master thesis.

First and foremost, I would like to thank my research supervisors, Professor Pierre Defourny and Professor Mathieu Javaux from the Earth and Life Institute, UCLouvain. A special thanks goes to Professor Mathieu Javaux for giving me the opportunity to address this interesting topic and for his always prompt answers/responses to my requests. Equally, my sincere appreciation goes to Professor Pierre Defourny for the relevant guidance, support and useful advice he provided me throughout this journey.

I am thankful to the researchers attached to my supervisors' laboratory (ENGE and GERU) for the time they devoted to support me with expert explanation.

I would also like to show gratitude to Doctor Julien Radoux and Damien Delforge, for having agreed to be members of my jury.

Without the support from the administrations of the Walloon Region and the Walloon Brabant Province (WBP) and especially their willingness to closely cooperate, I couldn't get insight in the procedures in use in those matters as well as in the essential data about the studied storm water basins. In particular, my appreciation goes to Alexandra Bauwens and Vincent Balthazar, members of the mapping and hydrology department of the WBP for giving me the access to the provincial platform of flood risk management. Similarly, I would like to thank Nicolas Feltz for allowing me to access the water-level time series collected by the Aqualim water-level network he is in charge of.

I would never be thankful enough to my parents for their never-ending patience and support, as well as the opportunity they gave me to complete this university master's degree.

Finally, I warmly thank my fellow members (Céline Champagne, Julien Herinckx, Dorian Lessire, Baptiste Deleuze and Alexis Del Marmol) for creating a highly motivating work atmosphere in the student workspace. I would like to end by thanking Thomas Huyberechts for his precious help and support throughout this thesis.

Table of Content

Introduction	1
Chapter I. Stormwater Basins	3
1. Definition of SB and Associated Typologies	3
2. Current Detection and Monitoring Methods of SB	5
2.1. Inventory	5
2.2. Monitoring	5
Chapter II. Literature Review	7
1. Remote Sensing Capabilities for Monitoring and Detection of SB	7
1.1. Detection of SB	7
1.2. Monitoring of Objects Similar to SB	8
1.3. Observations and Considerations for the Current Work	10
2. Platforms and Sensors used for Water Detection	10
3. Techniques for Water Identification	12
3.1. Threshold-based Methods	13
3.1.1. Spectral Signature of Water	13
3.1.2. Single-band Thresholding Method	13
3.1.3. Ratios and Water indices	14
3.1.4. Thresholding Limitations	14
3.1.5. Performance Comparison of Indices	15
3.1.6. Conclusion	15
3.2. Classifiers-based Methods	16
3.2.1. Un-supervised and Supervised Classification Methods	16
3.2.2. Parametric or Non-parametric Classification Methods	17
3.2.3. Pixel-based or Object-based Classification Methods	17
3.2.4. Common Classifiers	18
3.2.5. Conclusion	19
3.3. Combination of Threshold and Classifier-based Methods	20
4. Accuracy Assessment	20
5. Water Monitoring	21
6. Conclusion	23
Chapter III. Objectives	24
Chapter IV. Study Area	25
Chapter V. Material	29
1. Ground-Collected Data	29
1.1. SB Reference Dataset	29
1.2. Rainfall Time Series	30
1.3. Time Series of Water Levels Inside SB	30
2. Remote Sensing Data	30

2.1.	Orthophotoplans	30
2.2.	Sentinel-2 Image Times Series	30
2.2.1.	Description of Sentinel-2 Images	30
2.2.2.	Image Acquisition	31
2.2.3.	Atmospheric Correction	31
2.2.4.	Clouds and Missing Values Removal	32
2.2.5.	Water Indices Calculation	32
2.2.6.	Creation of Calibration and Validation Images	33
3.	Softwares	34
Chapter VI. Methods		35
1.	Overall Approach	35
2.	SB Dataset Preparation and Characterization of their Detectability	36
2.1.	Delineation of SB	36
2.2.	Detectability of SB	37
2.2.1.	Assignment of Detectability Level by Photointerpretation	37
2.2.2.	Validation of the Photointerpreation	38
3.	Calibration and Validation of the Model for Water Detection inside SB	39
3.1.	Labelling SB with their corresponding bottom type (Reference values)	39
3.2.	Training and Calibration of the Classification Model on Excellent Basins	40
3.2.1.	Training	40
3.2.2.	Calibration	40
3.2.3.	Determination of the Most Effective Subset of Features	41
3.3.	Validation of the Classification Model	43
4.	Water Monitoring inside SB	43
4.1.	Application of the Classification Model on TS for both medium and Excellent SB	43
4.2.	Validation of the Temporal Evolutions	44
Chapter VII. Results		45
1.	SB Dataset Characterization	45
1.1.	Delineation of SB	45
1.2.	Detectability of SB	45
1.2.1.	Validity Assessment Outcomes	46
1.2.2.	SB Detectability Levels	48
1.3.	Conclusion	49
2.	Calibration and Validation of the Classification Model on Excellent SB	50
2.1.	Bottom Types Labels	50
2.2.	Calibration of the Classification Model	52
2.2.1.	Determination of the Subset of Features	52
2.2.2.	Calibration of the Model	57
2.2.3.	Conclusion	61
2.3.	Validation of the Model over Excellent and Medium SB	61
2.3.1.	Homogeneous Excellent SB	62

2.3.2.	Heterogeneous Excellent SB	62
2.3.3.	Medium SB	63
2.3.4.	Conclusion	65
3.	Monitoring of Water Inside SB	66
3.1.	Application of the Model over the Time Series for both Medium and Excellent SB	66
3.1.1.	Time Series of Classified Sentinel-2 Images : Strengths & Weaknesses	66
3.1.2.	Reduction of Frequency of Observations by Clouds	68
3.1.3.	Observed Evolutions	69
3.2.	Validation of the Temporal Evolutions	72
Chapter VIII. Discussion		74
1.	Discussion of SB Detectability	74
2.	Discussion of the Model Calibration and Validation	76
3.	Discussion of SB Monitoring	76
4.	Discussion of the Method	77
5.	Perspectives	78
Conclusion		79
Bibliography		81
Webography		92
Appendix A. Illustrations of SB		I
Appendix B. List of non-delineated SB		IV
Appendix C. Separability Analysis		VI
1.	SFFS Algorithm and JM Distance	VI
2.	With Gini Index	VII
Appendix D. Accuracy Assessment of the Classification Model		IX
1.	Calibration of the Number of Training Areas	IX
2.	Calibration of the Number of Features and Images	IX
3.	Validity of the Model for Medium and Excellent with Reference Images	X
Appendix E. Accuracy Assessment of the Temporal Evolution of Water Proportion for Medium and Excellent SB		XI
1.	Illustration of water proportion's evolution	XI
2.	Algae Bloom	XI
3.	Validation Results	XII

List of Figures

1	Stormwater basins appearances	4
2	Absorption characteristics of pure water.	13
3	Study area extent	25
4	Sub-watersheds and principal watercourses of Walloon Brabant province and their corresponding elevation	26
5	Common locations of stormwater basins	27
6	Categorization and corresponding administrators of the water system in Belgium (Walloon Brabant Province, 2019)	27
7	Workflow of clouds and missing values removal	32
8	Workflow of overall methodology established to monitor the temporal variation of water presence inside observable SB	35
9	Controversial challenging (a,b) and medium (c) SB not upgraded to the higher level .	47
10	Spatial distribution of detectability levels assigned to SB	48
11	Illustration of the detectability levels of SB (challenging, medium and excellent) using orthophotoplans and Sentinel-2 to present the visibility and the size/shape of SB . . .	49
12	Bottom type discrimination using Sentinel-2 false-color composite	50
13	Discrimination between geomembrane and water bottom types with false color, true color and Band 2 (Blue), NDWI and AWEIsh Sentinel-2 composites	51
14	Mean spectral signature of the bottom types for the explanatory variables	52
15	Density distributions of bottom types' reflectance for poor, medium and excellent discriminating features	53
16	Visualization of the discriminating power for pairs of features	53
17	Correlation among the Sentinel-2 bands and retained water indices	55
18	Temporal variability of the Jeffries-Matushita distance of each feature	56
19	Jeffries-Matushita distance according to the number of features selected by the SFFS algorithm	56
20	Comparison of the classification products obtained by training the model with training areas of diverse number and nature	60
21	Overall accuracy of the Random Forest classification models depending on the reference image and the number of features selected for their training	61
22	Ability of the model to depict the correct bottom types inside the homogeneous excellent, except in presence of shadows	62
23	Ability of the model to depict the heterogeneity inside the heterogeneous excellent SB	63
24	Ability of the model to depict the correct bottom types inside the medium, except in presence of shadows	64
26	Ability of the model to depict the temporal evolution of water proportion in dynamic excellent SB	66
27	Ability of the model to depict the temporal evolution of water proportion in steady excellent SB	66
28	Overestimation of water proportion resulting from the winter conditions (snow cover, high soil moisture or low sun's course) of Sentinel-2 images	67
29	Underestimation of the water proportion caused by the residual geometric error . . .	68

30	Underestimation of water proportion caused by the presence of algae at the SB surface	68
31	Types of temporal evolution of proportion observed in the SB	69
32	Diversity of SB's state across a part of the communes of Louvain-la-Neuve and Chaumont-Gistoux.	70
33	Diversity of SB states on the N25 over Genappe.	70
34	Illustration of the diversity of water temporal evolution encountered inside the SB depicted in Figures 32 and 33.	71
35	Validation of the obtained temporal evolution with daily water heights.	73
A. 1	Illustrations of excellent SB	I
A. 2	Illustrations of medium SB	II
A. 3	Illustrations of challenging SB	III
C. 1	Feature importance ranking depending on their decrease effect on the Gini index for water discrimination	VIII
E. 1	Ability of the model to depict the temporal evolution of water proportion in dynamic excellent SB	XI
E. 2	Underestimation of water proportion caused by the presence of algae at the SB surface	XI

List of Tables

1	Characteristics of remote sensors for water surface detection	11
2	Spectral and spatial resolution of the Sentinel-2 bands	31
3	Definition of the selected water indices	33
4	Definition of the detectability levels (challenging, medium and excellent) for the size/shape variable and the visibility variable. *After the application of a negative buffer of 10m.	38
5	Number of samples used for classification of medium and excellent SB types in each tile UES or UFS	44
6	Accuracy of the photointerpretation (before revaluation)	46
7	Accuracy of the photointerpretation (after revaluation)	47
C. 1	Best subsets of features obtained for the different dates by letting the algorithm freely choose among the complete set of features.	VI
C. 2	Best subsets of features obtained for the different dates by letting the algorithm freely choose among the following reduced set of features :Band 2, Band 4, Band 8A, Band 11, NDWI, AWEInsh, AWEIsh, LSWI and s.	VI
C. 3	Best subsets of features obtained for the different dates by letting the algorithm freely choose among the following reduced set of features (Band 2, Band 4, Band 8A, Band 11, NDWI, AWEInsh, AWEIsh, LSWI and s) to which the Band 8 has been added	VI
C. 4	Best subsets of features obtained for the different dates by letting the algorithm freely choose among the following reduced set of features (Band 2, Band 4, Band 8A, Band 11, NDWI, AWEInsh, AWEIsh, LSWI and s) to which the Bands 8 and 7 have been added	VI
C. 5	Best subsets of features obtained for the different dates by letting the algorithm freely choose among the following reduced set of features (Band 2, Band 4, Band 8A, Band 11, NDWI, AWEInsh, AWEIsh, LSWI and s) to which the Bands 8 and 7, as well as the indice mNDWI have been added	VI
D. 1	Confusion matrices of the eight RF classification models developed to calibrate the number of training areas.	IX
D. 2	Confusion matrices of the eight RF classification models developed to calibrate the number of features and number of images.	IX
D. 3	Performance assessment of the model trained with mono-date versus multi-date	X
D. 4	Accuracy obtained when applying the retained model (Trained with 24 homogeneous SB training areas on the September image with a subset of features to predict 2 categories) on the three reference images for excellent and medium SB, with traditional and modal comparison.	X
E. 1	Validation performed on 10 images for all the SB covered by these images	XII
E. 2	Validation performed on 10 SB (Excellent) for all the available images	XII
E. 3	Validation performed on 10 SB (Medium) for all the available images	XII

List of Equations

2 Producer accuracy/Precision 20

3 User accuracy/Recall 20

4 F1-score 20

5 Quantity disagreement 21

6 Allocation disagreement 21

9 Jeffries-Matushita Distance 42

List of Abbreviations

ANN Artificial Neural Networks

ANOVA Analysis of variance

ASTER Advanced Spaceborne Thermal Emission and Reflection Radiometer

AWEInsh Automated Water Extraction Index for image with no shadow

AWEIsh Automated Water Extraction Index for image with shadows

CART Classification And Regression Tree

DT Decision Tree

ELIe Earth & Life Institute – environmental sciences

ENVISAT ENVIronment SATellite

EO Earth observation

ESA European Space Agency

GIS Geographic Information System

GPS Ground Positioning System

HSV Hue Saturation Value

ISODATA Iterative Self-Organizing Data Analysis

JM distance Jeffries-Matushita distance

LiDAR Light Detection And Ranging

LSWI Land surface water index

MACCS Multi-sensor Atmospheric Correction and Cloud Screening

MD Minimum Distance

MhD Mahalanobis Distance

mLSWI modified Land Surface Water Index

ML Machine Learning

mNDWI modified Normalized Difference Water Index

MODIS Moderate-Resolution Imaging Spectroradiometer

MSI Moisture Stress Index

MSI Multi-Spectral Instrument

MXL Maximum Likelihood classifier

NDVI Normalized Difference Vegetation Index

NDWI Normalized Difference Water Index

NN Nearest Neighbours

OBIA Object-based image analysis

PLUIES Prévention et Lutte contre les Inondations et leurs Effets sur les Sinistrés

PP Parallelepiped

PSF Point Spread Function

RBF Radial based Function

RF Random Forests

RS Remote Sensing

S2ITS Sentinel-2 Image Time Series

SAM Spectral Angle Mapper

SAR Synthetic Aperture Radar

SB Stormwater Basins

Sen2Agri Sentinel-2 for Agriculture

SFFS Sequential Forward Floating Selection

SPOT Système Probatoire pour l’Observation de la Terre

STIS Time Series of Satellite Images

SVM Support Vector Machines

SWM Sentinel Water Mask

TCW Tasselled Cap Wetness

TIFF Tagged Image File Format

UCLouvain Catholic University of Louvain-la-Neuve

UTM Universal Transverse Mercator

WBDA Water Bodies Detection Algorithm

WBP Walloon Brabant Province

WGS84 World Geodesic System reviewed in 1984

Introduction

Scientific and public attention for environmental issues and their potential negative effects are constantly growing in the recent years. It has steadily generated concerns on how we could predict, prevent and mitigate the impact they can have on our societies.

In that regard, water management is and remains a critical domain where action is definitively becoming a vital necessity. Because of the damages caused by massive flash floods or at the contrary due to its long-lasting scarcity, water has the propensity to affect populations and societies sometimes leading to dramatic humanitarian crises. It is therefore our collective responsibility to tackle these issues.

The construction of stormwater basins (SB) was one of the first measures implemented in the 1970s to reduce flood peaks and prevent downstream material and human damages. Since then, SB have continued to play a key role in flood management. However, the currently used SB inventory and monitoring methods remain insufficient to allow the authorities in charge to ensure any effective large-scale management. Innovative approaches could provide them the ability to decide and implement effective management actions. They therefore require improvement, for instance, through a greater use of new or disruptive technologies.

The situation described here before is also valid for Wallonia. The methods and techniques currently used by the Walloon administration do not provide a comprehensive understanding of the location and the operational state of SB. There is a need at both regional and provincial scale as the current approach does not allow any integrated and coordinated management.

In this context, the province of Walloon Brabant mandated in 2012 the Earth Life Institute – environmental sciences (ELIe) to conduct an unprecedented centralization of the existing data related to SB. The data were collected by means of questionnaires and were integrated into a provincial platform to manage floods risks through a map of SB. This initiative is a first step in the right direction but is clearly not sufficient to properly respond to the current and future necessities. There is a clear need to develop new methods for gathering and managing the much-needed information.

Remote sensing constitutes in that regard an unexploited and promising method. Unfortunately, during the past years, this option remained justifiably overlooked. This is mainly due, on the one hand, to the relatively high costs of acquiring satellite images and, on the other hand, the ‘poor’ suitability level of sensors with regard to spatial coverage and spatial and temporal resolutions. The recent launch of ESA’s Sentinel-2 satellites has brought new perspectives for any development of an improved regional and provincial scale SB monitoring.

The following research study builds on that renewed perspective while at the same time offers a continuity to the earlier initial project launched by the Walloon Brabant Province. The study aims at developing realistic innovative ways of acquiring the required information about SB. In other words, the overall objective of this thesis is to evaluate the feasibility of monitoring the hydrological functioning of SB using Sentinel-2 image time series over the province of Walloon Brabant.

The greatest interest of this work lies in the fact that it breaks new grounds. However in comparison to other similar water-related research, the challenges to be expected here relate to specifically to SB : the diversity of the SB water characteristics, the poor visibility of certain SB due to obscuring land-

scape elements, the diversity of SB's appearance (size/shape, bottom type) as well as the resemblance of SB's with other landscape features (crops, lakes, roads, etc.). The fact that SB are rather small, combined with quite a spread spatial spacing over the study area, constitutes an additional challenge to be tackled.

The current master thesis addresses the subject through a structured analysis along eight chapters.

The first one provides, on the one hand, the necessary terminology with regard to SB definition and typologies. On the other hand, it gives the current state of the art with regard to stormwater basins detection and monitoring.

Further on, a literature review (Chapter two) provides a comparison between the current methods for SB inventory and monitoring and the possibilities offered by existing remote sensing capabilities and applications. The aim is to highlight an appropriate approach of SB detection and monitoring using remote sensing while also listing the limits and constraints the research will have to cope with. To complete that scene setter, this section also offers an overview of the available platforms and sensors. This serve to identify the most appropriate sources of images for the current work in terms of spatial, temporal, spectral resolutions as well as spatial coverage. A description of the methods used to depict spatio-temporal water change of SB out of the analysis of image time series is also performed. This concludes the chapter that globally supports the definition of the study objectives and the related methodological choice to achieve them.

As a result of this process, the third chapter further clarifies the thesis objectives. The overall aim is to assess the feasibility of monitoring the hydrological functioning of SB using Sentinel-2 time series. However, from the beginning, the factual premises of the study pushed to identify three more specific objectives : firstly, the SB dataset preparation and the ex-ante characterization of the SB potential detectability based on the adequacy between the SB's characteristics and the Sentinel-2 images limitations ; secondly, the establishment of a classification model that exploit the information provided by three Sentinel-2 calibration/validation images in order to detect the presence of water inside the observable SB and thirdly, the potential of the Sentinel-2 time series along with the developed classification model to depict temporal variation of the surface of such observable SB.

The fourth chapter sets the scene further and gives a brief presentation of the Walloon Brabant Province, which constitutes the geographical area that serves as a study case.

The fifth chapter lists from a material perspective, the ground-collected and remote sensing data sources, as well as the softwares which served to generate maps and achieve automation of various operations.

Chapter six describes the methodology implemented to assess the feasibility of using Sentinel-2 images to monitor SB while the chapter seven tackles the essence of the study as it is dedicated to the description and the display of the different experiment's results.

The eighth chapter first addresses the interpretation and the discussion in light of the results obtained through the previously completed research steps. It continues with a brief commented recall of the methodology's limitations and logically introduces some future prospects.

The final part offers overall final conclusions of this study, covering all previously addressed perspectives and results.

Chapter I. Stormwater Basins

The following section introduces the subject of this thesis firstly, by defining stormwater basins (SB) and their associated typologies and secondly, by describing how the SB identification and monitoring is presently being performed.

1. Definition of SB and Associated Typologies

As the envisaged research combine complex interconnected subjects like hydrology and remote sensing observation techniques, it seems very much needed to clarify the necessary definitions and typology to be used for the rest of the work, before tackling the effective subject's thesis.

The core of this thesis relates to stormwater basins (SB). SB are defined as man-made hydraulic structures built to mitigate the impact of urban stormwater runoff. This impact can be related to both quantitative and qualitative aspects of running-off stormwater. On the one hand, SB alleviate the quantitative impact of stormwater with its capacity to store temporarily stormwater and to further prevent downstream flooding and erosion. On the other hand, SB deal with the qualitative aspect of stormwater by treating the collected water before being released in the natural environment. The SB addressed in this study collect rainwater exclusively. Therefore, they have to be distinguished from basins located on a single sewer system, which store a mix of waste and rainwater before their treatment into a waste water treatment plant. The SB of interest can be connected to the pluvial network of a separate sewer system only (not unitary sewer system), or off-grid. In addition, this thesis considers only the open-air basins since the buried ones can not be detected with remote sensing technology. From a morphological perspective, those basins can be defined as impoundment areas characterized by embankments or excavations in the ground. Moreover, as they are supposed to collect rainwater, they are often located in the collection zone of the watershed, besides roads and/or water courses. Their visual appearance is highly diverse (Figure 1). They often present a geometrical shape, as they are man-made structure. Their embankments can be covered with a geomembrane (to ensure waterproofing in case the soil is permeable) (Poulain *et al.*, 2011), made of concrete which can progressively be colonized by vegetation. Their typical size is ranging from hundreds to thousands of square meters.

A distinction is usually done among those basins based on their most frequent elementary state of the basin. Two types of basins are distinguished : wet and dry basins. By definition, wet basins always exhibit a water surface. The water level fluctuates according to precipitations and runoffs from the contributing areas. The storage capacity is limited by the tidal range of the lake, i.e. difference between the minimum and the maximum level of water (Bruxelles Environment, 2010b). In absence of rain, the basin may also be supplied by groundwater. In any cases, a sufficient depth is required to prevent the invasion of aquatic plants rising for the bottom of the basin and it discourages resuspension by keeping deposited sediments at the bottom (Bergue and Ruperd, 1994). According to this definition, the wet type of basins comprises run of river basins. In contrast to wet basins, dry basins are always empty except during or right after rain or snow melt (Laramie County Conservation District, 2016). The whole volume of the basin is dedicated to the storage of stormwater. As for the wet basins, the stored volume is released through an outlet pipe (towards the pluvial network or wells), via evaporation or evapotranspiration. In addition, when the infiltration is allowed, both bottom and banks of dry basins can infiltrate the incoming water. This represents a greater infiltration surface in dry basins than in

wet basins, where infiltration is often limited to the banks.

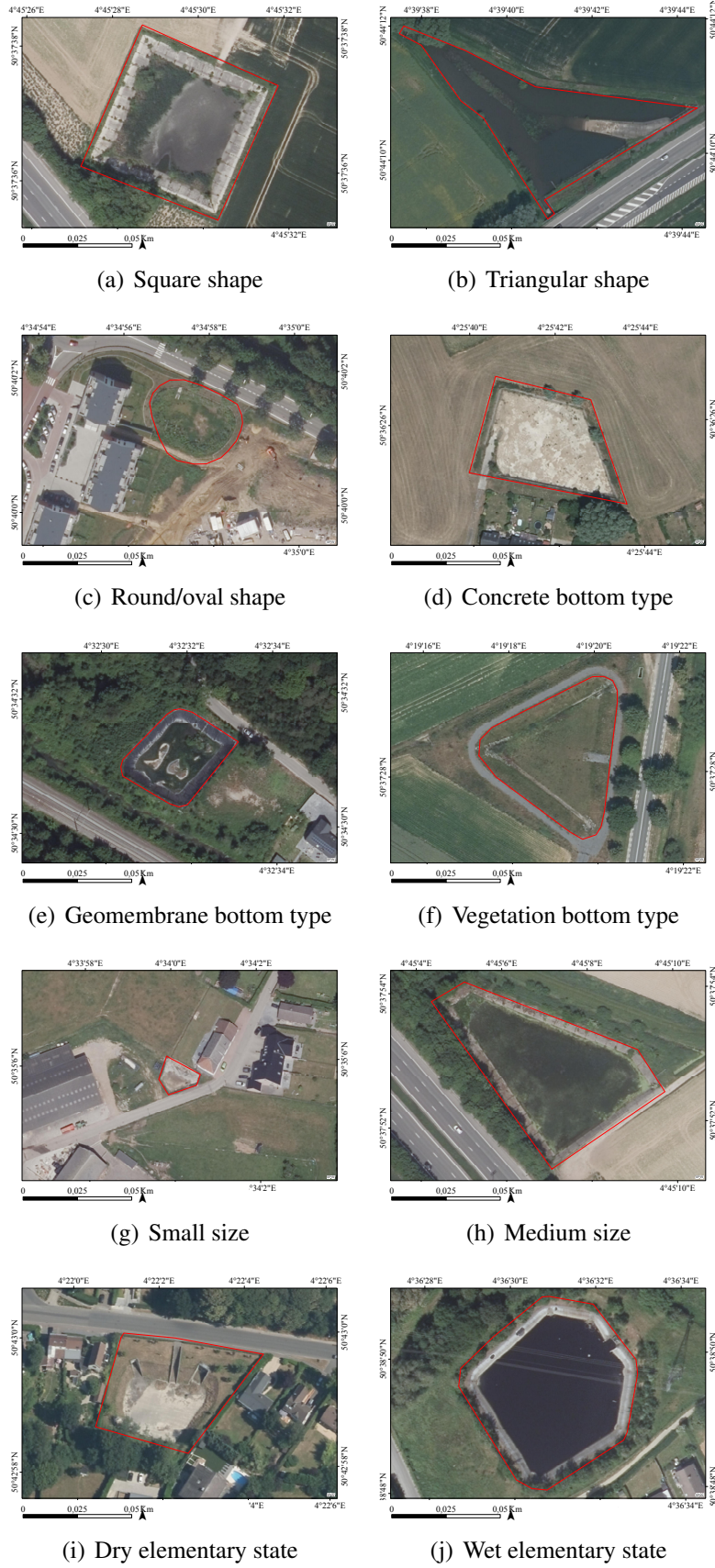


FIGURE 1 – Stormwater basins display various appearances in terms of shape (a-b-c), bottom type (d-e-f), size (g-h) and elementary state (i-j). To be compared, images are displayed at the same spatial scale (1 : 650)

In the literature, many combinations of terms refer to stormwater basins due to existence and use of synonyms. The term « ponds », for example, is indifferently used to refer to basins. In addition, wet and dry basins are also referred to as retention/impoundment and detention basins respectively. For the sake of simplification, the thesis will distinguish the basin according to their type (wet or dry) when it is crucial for the understanding and will comprise both of them under the term ‘stormwater basin’ (SB) otherwise.

2. Current Detection and Monitoring Methods of SB

2.1. Inventory

The inventory of the SB (geographical location, characteristics, etc.) of a country or a region is for the time being usually done in a decentralized way under the lead of the competent local authorities mainly at provincial and communal level. The local inventories remain thus limited to the area of responsibilities of those authorities. Furthermore, the collection of the relevant data is essentially done through a registration process during the design phase, when the necessary construction’s authorization needs to be granted. The so collected dataset remains therefore only based on initial design and construction data and is afterwards rarely updated. When the dataset needs to be refreshed, it is usually done through the launch of a separate specific process reaching out to all potential involved actors (authorities and subject matter experts) to get from them the necessary updated information.

Belgium is in that regard no exception. Fully aware that the Belgian institutional organization is far from being simple and ideal from that perspective, the Belgian provincial authorities of Brabant Walloon have conducted an unprecedented centralization of the existing data. More specifically, the province of Walloon Brabant mandated the Earth & Life Institute – environmental sciences (ELIe) to develop a provincial platform to manage floods risks, comprising a comprehensive map containing the position of all Walloon SB. This project has been conducted, on the one hand, through the collection of all available datasets and, on the other hand, by means of questionnaires and interviews to complement the lack of available data. The same methodology has been observed in a Swedish study (Starzec *et al.*, 2005).

Both cases highlight the limits and constraints of the methodology. This registration approach is labour intensive and, at the same time, relatively uncertain with regard to the accuracy and reliability of the provided data. The quality of the provided data effectively depends on the rigor shown by the involved administrative staff and their effective level of commitment to assist with this type of duty. Contributing to such a broader overarching task as a complex national or regional inventory is rarely welcome by any administration, mainly because it is a very demanding job, quite laborious. In the meantime, it generally exceeds their formal governance responsibility.

Innovative approaches and capabilities avoiding those pitfalls would be very beneficial for all the concerned actors.

While the registration, localisation and identification of SB is obviously challenging, ensuring the well-functioning monitoring of the of those basins still needs to be investigated.

2.2. Monitoring

Practically, as it is the case for SB’s inventory, their monitoring also remains, for the time being, essentially based on manual methods. The SB monitoring tasks are basically realized through human

intervention and visual inspections conducted by the managers and staff in charge of the related infrastructure. Such visual inspections and *in situ* visits are characterized by low replicability in time and space. The measurements made at those occasions allow some basic assessment of the SB state regarding its retention capacity (level of silting, quantity of retained water) but are too sparse to properly report on the hydrological functioning (filling/drain) of the infrastructure. Furthermore, performing those inspections on several or all SB in a single region is time-consuming, laborious and costly, especially knowing the number of SB to consider, their spread spatial spacing and in most of the cases their difficult accessibility.

To mitigate those constraints, an alternative consists of equipping some SB with sensors and webcams allowing a continuous monitoring. Sensors probes serve to measure flow rates, water depth, pluviometry and changes in sediment height, whereas webcams can be used to film water height changes.

When correctly calibrated and in operable conditions, sensors probes and cameras constitute an accurate means of assessing the hydrological functioning of SB. Though already providing a more effective method, similarly to visual inspection, continuous *in situ* monitoring remains time-consuming and costly because it also additionally requires the physical installation of the sensors, the integration of data in a simulation model as well as their analysis and interpretation afterwards. Moreover, the main drawback of those traditional methods is to cover a very limited spatial scope as the related monitoring is only applicable at one SB at a time.

As aforementioned, the classical traditional methods have shown limits and constraints and can quite easily be qualified as inefficient for the study of the location and functioning of SB at a regional scale. At the contrary, Remote Sensing (RS) or Earth Observation (EO) techniques and capabilities seem to present interesting advantages, which are also currently in constant evolution.

The first noticeable advantage is the possibility to get access to the required data for free what interestingly reduced the overall financial costs. Furthermore, the theoretical frequency of RS data acquisition is often higher than traditional methods. Consequently, properly exploiting the opportunities and capabilities EO offers, should a priori allow a time resolution (temporal resolution) varying between the one ensured by continuous sensor-monitoring and the low one provided through punctual visual inspections. In addition, remote sensing data currently offer a sufficient spatial coverage to allow the assessment of a number of SB at the same time. This may support the idea that RS shall solve the issue of low replicability of measurement. However, the spatial resolution of currently freely accessible satellite data which offer a sufficient spatial coverage to allow the simultaneous assessment of all SB, will in any case remain a constraint that may not be ignored. From that perspective, the sensor performance will eventually determine the minimal size under which a SB will not be detected.

To conclude, this section illustrated, on the one hand, the inefficiency of the traditional methods of SB inventory for the study of the functioning of SB at a regional scale. On the other hand, it revealed that remote sensing seems, at this stage and based on all the aforementioned considerations, to be a very interesting alternative to the more traditional methods because of its larger spatial coverage than continuous monitoring and its higher frequency of observation than punctual visual inspections. The further aim of this work is now to study more deeply the remote sensing capabilities in regard of SB monitoring and management control tasks.

Chapter II. Literature Review

This chapter aims to provide a state-of-the-art synthesis of the key concepts for this thesis and is structured as follows. This literature review starts with a comparison of the state of play of the current detection and monitoring methods (Chapter I) with the possibilities that existing remote sensing's capabilities and applications could offer (Section 1). This should highlight an appropriate approach of SB detection and monitoring using remote sensing while, at the same time, allowing the identification of some limits and constraints to consider in that regard for the final study. The second section gives an overview of the platforms and sensors usually used for water detection. This helps to identify the most appropriate source of images for the current work in terms of spatial, temporal, spectral resolutions and spatial coverage. A description of the methods used to depict water change of SB and (their corresponding accuracy) out of the analysis of image time series is also performed (Sections 3,4 and 5). The layout of the sections forming this literature review is inspired by the informative review of Huang *et al.*(2018). The overall approach globally supports the definition of the objectives and related methodological choices to achieve them (Section 6).

1. Remote Sensing Capabilities for Monitoring and Detection of SB

Most of the existing literature dealing directly or indirectly with SBs has generally been focusing on the assessment of their functioning performances. This entails, amongst others, the assessment of their pollutant removal ability or their capacity to prevent flooding and erosion peaks (Ivanovsky *et al.*, 2018). Surprisingly, the use of remote sensing techniques and data is virtually non-existent in the current literature.

1.1. Detection of SB

To date, only two researchers Liu and Wang (2008) used remote sensing data in addressing the issue of detecting SBs. Thanks to the Light Detection And Ranging (LiDAR) elevation model, they succeeded in introducing an innovative detection method. In a first stage, Liu and Wang (2006) developed a surface depression delineating method to mark all surface depressions in the LiDAR elevation data. In a later stage, they performed a classification of the surface depressions on the basis of their geometric characteristics including the area, depth, compactness and shape (Liu and Wang, 2008). Each patch of the surface depressions was treated as a single spatial object, created by applying the blob-colouring algorithm (Ballard and Brown, 1982) onto the surface depressions data and the geometric measures of the surface objects were then calculated in a geographic information system (GIS). Finally, the machine learning algorithm C5 (Quinlan, 1999) built the rule for differentiating the detention basins from other objects (Liu and Wang, 2008). The location of the detected basins was eventually confirmed by a visual inspection, before being incorporated into the map of SBs. This method has proven to be more efficient and reliable than traditional field-based levelling surveys, and it avoids the arbitrariness of human intervention. Moreover, the use of remote sensing technology enables cost-effective data collection over large areas.

The Public Service of Wallonia (SPW) also relied on a LiDAR elevation model to conduct a flood-related project. This model enabled the mapping of flood conveyance routes (LIDAXES), which helped to identify the areas at risk of flooding or mudflows caused by the stormwater runoff. However, so far, there have been no plan to integrate SB in this project, nor to use their product for the detection

of SB.

1.2. Monitoring of Objects Similar to SB

Due to the lack of specific literature related in particular to SB remote sensing, it seems to be no other alternative than to exploit at best the other numerous existing similar research though from a broader perspective. The aim would then be to explore a substitute approach and to investigate which RS techniques and methodologies that have already been used to detect and monitor relevant objects that share similarities with SB.

From that perspective, aquaculture ponds, water bodies and flooding areas have to be considered. Comparing their respective interesting similarities and differences with SB should eventually allow the identification of an appropriate substitute approach.

Aquaculture ponds are of particular interest since they are characterized by a water surface surrounded by pond embankments and their assembly often presents a geometrical shape, just like SB. Several relevant papers have focussed the use of multispectral and radar imagery for mapping aquaculture facilities in China (Zhang *et al.*, 2010), in India (Sridhar *et al.*, 2008; Jayanthi 2011 and Dwivedi and Kandrika, 2005), in Thailand (Travagli *et al.*, 2004), in the Philippines (Travaglia *et al.*, 2004) and in Vietnam (Virdis, 2013). All of them were performing object-based classification. Some studies related to aquaculture ponds exploit similarities with SB and therefore are in some way transposable. It is the case of Zhang (2013) since its methodology, developed with a region growing segmentation integrated edge detection, is based on the recognition of edges, a similar key feature of SB. Another study makes use of four indicators to determine the actual presence of aquaculture ponds : water presence, existence of a bridge, an aerator and absence of vegetation (Gusmawati, 2017). In this study, replicability is limited to the only common element, i.e. the water presence.

On the other hand, with regard to differences, aquaculture ponds differ from SB when considering the water presence, their organizational juxtaposition and the limited presence of confounding elements like vegetation. Although the element water surface is the only common element to both objects, the presence of water in the case of aquaculture ponds is always permanent, unless the ponds are inactive. In the case of SB the presence of water is usually rather dynamic, subject to variations, for instance according to rainfalls.

Water bodies are landscape elements which also resemble SB because of their water surface. However, the fact that they generally have non-geometrical shape, makes them different. Numerous studies have been conducted to detect waterbodies (Wang and Hie, 2018), while lots of them were focussed on their detection and monitoring especially in arid areas as water presence in these regions is obviously very important. This subject therefore attracts scientific attention. Some studies analyse their geographical repartition in relation with disease transmission. Others investigate their evolution in order to assess the potential impact they can have on resources' availability (Soti, 2010). All those studies have in common the aim to detect water and, therefore, present a relevant interest for the current work. However, these cases are quite specific and, from the sole detection point of view, significantly differ from SB water surface in temperate regions. Because of the relative scarcity of these waterbodies as of the arid climate, the observed water surface offers a strong contrast with its surrounding environment. This is absolutely not the case in regions with temperate climate where water surface can be easily confused with shadows and where the observation can be obstructed by existing vegetation. Water body studies within a similar climate would be more interesting as they would include the difficulty

to differentiate water surface from existing vegetation as well as from shadow effects.

In this vein, d'Andrimont *et al.* (2017) have recently done extensive research on waterbodies detection in Wallonia. This study is of great interest as it simultaneously covers the same two key aspects to be analysed in the current work. On the one hand, d'Andrimont *et al.* (2017) covered the same geographical area of interest as the one identified for this work, and, on the other hand, they also specifically studied the detection of the same objects, namely water bodies including SB. This work constitutes a very interesting source of guidance, firstly because the detection of the Belgian and Walloon water bodies has been made through the use of high resolution images and, secondly because of the useful unique experience it represents with regard to the level of effort it has mobilised both in terms of searcher's workload and in terms of computing power.

More specifically, d'Andrimont *et al.* (2017) proposed a framework to update hyperspatial resolution (<1 m) land thematic mapping over large areas by handling multi-source and heterogeneous data. A choice was made to use this framework to delineate surface water bodies in the southern part of Belgium (17,000 km²). Observations were collected by different platforms and sensor types (25-cm passive optical and 1-m LiDAR) as well as different instruments (three cameras and two LiDARs) used in heterogeneous observation conditions (date, time, and sun angle). Also the chosen targets showed a great degree of diversity, on the one hand, with regard to the water characteristics (morphology, origin, composition) and, on the other hand, with regard to the neighbourhood effects (shadow). The proposed framework dealt with the heterogeneity from both observations and targets, aiming at efficiently delineating water bodies. As a result, the proposed framework detects more than 98% of water bodies greater than 100m², with an edge accuracy below 1-m over large areas. The efficiency of detection decreases to 48,15% for smaller water bodies, with a size between 50 and 100 m², with a 1-m delineation.

In complement but still in relation with our approach with regard to waterbodies, another interesting work done by Pekel *et al.* (2016) have served to map all waterbodies of the world with a resolution of 30 meters. The datasets generated during the current study are available from <https://global-surface-water.appspot.com>. Although this spatial resolution is very high for such a large area, a 30 m resolution remains too coarse to enable SB detection at local/regional level. To achieve such a global scale coverage, Pekel *et al.* (2013) had previously developed, demonstrated and validated the adequacy of a generic multi-temporal and multi-spectral image analysis method to detect water surfaces automatically, and to monitor them in near real time at the African continental scale. The proposed approach was based on a transformation of the RGB color space into Hue Saturation Value (HSV) using MODIS multi-spectral time series data.

Finally, flooding areas demonstrate interesting forms of similarity with SB. Studies on flooding areas (Jain *et al.*, 2005; Malinowski *et al.*, 2015) offer different advantages. Firstly, they are generally dimensioned to manage the dynamic character of water presence. Secondly, compared to studies on waterbodies, these works also take into account the presence of additional disturbing elements like shadows, vegetation or clouds (Lin *et al.*, 2016). The only difference of the flooding areas with SB is that stormwater is generally concentrated in geometrical shape reservoir, what will more rarely be the case by natural flooding. This last point put aside, the study case of flooding areas presents, above the common criteria of water presence, additional interesting features to be considered.

1.3. Observations and Considerations for the Current Work

All in all, the current overview of the available studies allowed interesting observations and considerations which will help choices to be made for this work, in particular, regarding which RS techniques to be further explored, on the one hand, and the level of resolution to favour, on the other hand. It also indicated that water presence is the main common element to be looked at when remote sensing is used for detection. It also drew the attention on other useful elements which could have significant influence on the feasibility as well as the quality of remote sensing observation.

Although the study of d'Andrimont *et al.* (2017) offers an interesting basis to work on, the current work cannot limit itself to reproduce the same study. Since Copernicus Sentinel satellites open new opportunities with high revisit time capabilities, it should rather look at exploring other RS possibilities to detect and monitor SB. Therefore, and based on the other aforementioned observations, it would be wise to try expanding the scope of the research beyond the LIDAR elevation model technique, the sole conclusive one in that matter, up to now. It could in the future again be useful to pursue with this technique but only once the expected new developments aiming at greater performances are confirmed and effective. This technique would, however, only serve the detection of SB as the LIDAR coverage is updated every two years (or more). In the meantime, the most ambitious choice would be to study how RS techniques and capabilities that were already used to detect and monitor objects that present relevant similarities with SB can serve the purpose. That is why the focus should be placed on the detection of the presence of water as the omnipresent element common to all those objects. Both the static/permanent as the dynamic presence of water should initially be considered.

The same overview allows the reader to note the key importance to favour medium (10-100m) to high resolution sensors (<5m). Once again, the work of d'Andrimont *et al.* (2017) which was executed with very high resolution images, has already provided interesting lessons on the necessary level of resolution to select, both in terms of research workload as well as computing power. It would be useless to focus again on the same parameters while opting for an intermediary level of resolution (0,25m and 30m) could interestingly break new ground though still allowing enough probabilities to detect normal-sized SB.

Expanding on these parameters and research options, the next section will proceed with an analysis of water extraction platforms and sensors to help identify the most suitable ones to detect and monitor SB.

2. Platforms and Sensors used for Water Detection

Platforms used for water body extraction are space-borne, air-borne or terrestrial. This section reviews only space-borne platforms (satellites) since the high temporal resolutions required for the purpose of this work are given by the remotest platforms. One can distinguish between two categories of sensors – passive sensors and active sensors.

Passive sensors measure mostly the reflected sunlight in the visible and the infrared spectral (d'Andrimont *et al.*, 2017; Nath and Deb, 2010). They are influenced by weather conditions since the short wavelength radiation they measure are easily stopped by obstacles like clouds or vegetation. This is especially problematic for monitoring of SB since SB filling occurs during rainy events. Nevertheless, optical satellite sensors have been widely used in water body extraction research due to high data availability, as well as suitable temporal and spatial resolution (Huang *et al.*, 2018). Active sensors emit

their own electromagnetic radiation towards the target and they record the timing and the intensity of the reflected radiation to draw valuable information about the land type (d'Andrimont *et al.*, 2017). Consequently, they no longer rely on the solar radiation which enables them to sense day and night regardless of the weather conditions. Another advantage is its ability to penetrate cloud coverage and a certain vegetation coverage due to the usage of long wavelength radiation (Huang *et al.*, 2018). Active optical Light Detection and Ranging (LiDAR) and active Synthetic Aperture Radar (SAR) are active sensors usually used to detect water.

Over time, several passive sensors have proved to be capable of mapping water surfaces. Huang *et al.* (2018) distinguished them according to their spatial resolution : coarse (<200m), moderate (10-100m) and high (<5m).

TABLE 1 – Characteristics of commonly used space-borne sensors for water surface detection (Source : Huang *et al.*, 2018) *The area coverage (swath) determines the scale of application : L, landscape; R, regional; G, global; L-R, landscape to regional; L-G, landscape to global; R-G, regional to global.

Sensor group	Satellite/sensor	Number of bands	Spatial resolution [m]	Temporal resolution [day]	Maximum swath width at nadir [km]	Scale of application*	Data distribution policy (costs)	Data availability
Coarse resolution sensor	NOAA/AVHRR	5	1 100	0,5	2 800	R-G	No	1978-
	MODIS	36	250-1 000	0,5	2 330	R-G	No	1999-
	Suomi NPP/VIIRS	22	375-750	0,5	3 040	R-G	No	2012-
	MERIS	15	300	3	1 150	R-G	No	2002-2012
	Sentinel-3/ OLCI	21	300	2	1 270	R-G	No	2016-
Medium resolution sensor	Landsat/MSS-TM-ETM ⁺ -OLI	4-9	15-18	16	185	L-G	No	1972-
	SPOT	4-5	2,5-20	26	120	L-R	Yes	1986-
	ASTER	14	15-90	16	60	L-G	No	1999-
	Sentinel-2/MSI	13	10-60	5	290	L-R	No	2015-
High resolution sensor	IKONOS	5	1-4	1,5-3	11,3	L-R	Yes	1999-
	QuickBird	5	0,61-2,24	2,7	16,5	L	Yes	2001-
	WorldView	4-17	0,31-2,40	1-4	17,6	L	Yes	2007-
	RapidEye	5	5	1-5,5	77	L-R	Yes	2008-
	ZY-3	4	2,1-5,8	5	50	L-R	Yes	2012-
	GF-1/GF-2	5	1-16	4-5	800	L-R	Yes	2013-

Typical sensors and their specific characteristics are reported in Table 1. Even though they slightly differ, common characteristics can be outlined.

Coarse resolution sensors are providing daily coverage of the globe, which is notably facilitating the monitoring of large water bodies : wetlands (Guo *et al.*, 2017; Mahé *et al.*, 2011); rivers (Wolski and Murray, 2008) and lakes (Swenson and Wahr, 2009) for global assessments (Prigent *et al.*, 2007; Papa *et al.*, 2010; Klein *et al.*, 2015). However, their low spatial resolution remains inadequate for small reservoirs detection. As a proof, a single MODIS pixel corresponds to 6,45 ha, making it impossible to discern a SB usually measuring thousands of m² (0,1 ha) (Ogilvie *et al.*, 2018).

In contrast, high resolution sensors (Table 1) capture images at high spatial resolution (< 5m) which enable the detection of small object. Application of water detection with high resolution sensors can be found in Zhang *et al.* (2014), Yao *et al.* (2015) and Malinowski *et al.* (2015). However, their reduced spatial coverage makes them less suitable for medium-sized catchments. Furthermore, high resolution images are also more susceptible to shadows in mountainous or urban regions which may prevent an efficient detection of water (Sawaya *et al.*, 2003). Finally, the financial costs to acquire the data and the computational cost to process them are prohibitive for operational mid or long-term monitoring using high resolution images (Ogilvie *et al.*, 2018).

Since coarse and high resolution sensors are not able to discern SB due to their low spatial resolu-

tion and insufficient spatial coverage to represent a significant part of the studied zone in one image respectively, the focus turns to medium spatial resolution sensors.

Medium resolution sensors provide images with a spatial resolution ranging from 10m to 100 m over a spatial cover varying between 60km to 290 km. Their revisit time amounts from 0,5 to 5 days. The category of medium sensors includes Landsat, Système Probatoire de l'Observation de la Terre (SPOT), Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). Those sensors have proved to be effective in detecting surface water bodies or flood (Fisher and Danaher, 2013 ; Davranche *et al.*, 2010 ; Sivanpillai and Miller, 2010). Recent sensor such as the Multi-Spectral Instrument (MSI) on board Sentinel-2 (launched in 2015) is ensuring more than the continuity of the aforementioned satellites. The two Sentinel-2 satellites capture images at a spatial resolution of 10m over the entire globe every 5 days, providing enhanced opportunities. Du *et al.* (2016) and Yang *et al.* (2017) have confirmed their ability to map surface water.

Sensors on board of Landsat, ASTER and Sentinel 2 are generally preferred over SPOT thanks to their free availability. Moreover, as illustrated in Table 1, all the characteristics (spatial, temporal and spectral resolution and spatial coverage) of Sentinel-2 sensors are more interesting in comparison with Landsat and ASTER sensors. This advantage is related to its recency, linked with technological advances. Their downsides are also in someway related to their recency. Existing studies are consistent in their conclusion and confirm the ability of detecting and monitoring water bodies from Sentinel-2 images (Zhou, 2017 ; Yang, 2017 ; Du *et al.*, 2016). The only case where Landsat sensor should be preferred over Sentinel-2 sensor, is for study requiring a historical view of the natural phenomenon. In this case, Landsat could provide images from 1972 at best (Irons, 2012).

Given the advantages that active sensors provide, a combination of both passive and active sensor could possibly lead to better results. Several studies (Haas and Ban, 2017 ; Kaplan and Avdan, 2018) indicated the possibility to combine data from Sentinel-2 (passive) and data from Sentinel-1 (active). SAR sensors could enable to overcome obscuration from vegetation and clouds, often problematic in passive sensing. Besides the opportunities of this combination, one should also consider the need to account for all the variability sources among images collected by sensors (d'Andrimont *et al.*, 2017).

In sum, passive sensors with medium spatial resolution like Sentinel-2 appear to be the most appropriate for water detection and monitoring at a regional scale. A possibility is also to combine them with active sensors while accounting for all the variability sources. Once the most appropriate sensors have been identified, the collected images have to be processed in order to extract the valuable information.

3. Techniques for Water Identification

This section aims at giving an overview of frequently used techniques for water extraction while mentioning their advantages, limitations and application conditions. Two types of techniques can be distinguished : threshold-based methods and classification methods. A combination of them is also considered.

3.1. Threshold-based Methods

3.1.1. Spectral Signature of Water

The foundation principle of spectral and threshold-based methods consists of taking advantage of the differences in reflectance signatures of water in comparison with other objects along the electromagnetic spectrum to distinguish water from other land cover types. Most of the water-related studies concentrate on the optical signal. Studies that use thermal microwave bands are not included in this review as they usually have coarse spatial resolution (Ma *et al.*, 2019).

The absorptance of water at the surface increases with wavelengths starting with a minimum absorptance at 420 nm (Figure 2) reaching its maximum (90%) in the infrared portion of the spectrum (McCloy, 2005). Pure water thus reflects about 10% or less in the blue and green range as the absorptance is weaker, a smaller percentage in the red and certainly no energy in the infrared range (Richards, 2013). However, the water of inland surface is rarely pure and reflectance increases with the chlorophyll and suspended soil sediment content (Ma *et al.*, 2019) and consequently, is variable in space and time.

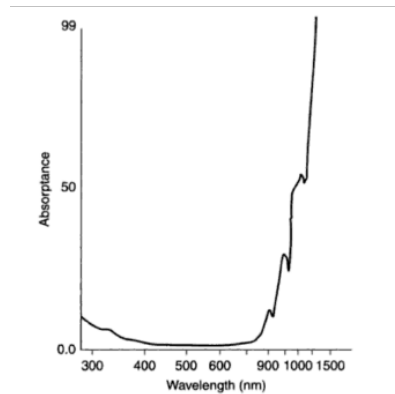


FIGURE 2 – Absorption characteristics of pure water. Most of the light is absorbed by water from the infrared part of the spectrum (McCloy (2005)).

3.1.2. Single-band Thresholding Method

The spectral and threshold-based approaches for water bodies mapping have undergone a succession of evolution. Single-band thresholding is the simplest method to derive a water map (Frazier and Page, 2000). The threshold value is derived from a manual histogram analysis of the single band images (Verpoorter *et al.*, 2012; Kallio *et al.* 2008; White and El Asmar, 1999; Smith, 1997; Overton, 1997; Manavalan *et al.*, 1993; Bennet, 1987) and the band used to extract water feature is often the infrared band (D’Andrimont *et al.*, 2017). Such simple method is easy and computationally not demanding but presents major limitations (Frazier and Page, 2000). The method demonstrates 1) an inefficient identification of mixed water pixels, 2) a confusion of water bodies with background noise and 3) variation of the threshold values according to the location and time of image acquisitions (Jiang *et al.*, 2014; Ji *et al.*, 2009). Consequently, the analysis cannot be restricted to one standard cut-off value (Jain *et al.*, 2005). Instead, thresholding computations should use a range of threshold values (Song *et al.*, 2002) and employ multiple band analysis to enhance accuracy (Verpoorter *et al.*, 2012).

3.1.3. Ratios and Water indices

To overcome the lack of stability of the threshold, bands were combined in ratio or in more complex combinations called indices. Spectral band ratios such as NIR/Red and Red/Green allowed for detection of water bodies (Puech, 1994; Lacaux *et al.*, 2007).

The first indices to be developed were not dedicated to water detection but still have proven to be efficient for this application. For instance, Crist (1984, 1985) put forward the Tasselled Cap Wetness index (TCW) derived from six bands of surface reflectance data and a threshold of 0 appears to be able to separate water and non-water objects. However, normalized ratios are often preferred since they reduce the impact of radiometric calibration error and remaining atmospheric perturbation. The Normalized Vegetation Difference Index (NDVI) was the first normalized index to be successfully used for water detection (Domenikiotis *et al.*, (2003)). Since its original aim is the detection of vegetation, it infer the presence of water from the absence of vegetation (negative values of NDVI) (Huang *et al.*, 2018). This way of detecting water presents a limited efficiency since negative values of NDVI also refers to other objects like snow, cloud, soil and built-up features. In response to these NDVI's limitations, McFeeters (1996) suggested the Normalized Difference Water Index (NDWI) which reverses the variables in the numerator and denominator of NDVI and uses the green band instead of the red band. In doing so, McFeeters (1996) utilizes the features of the relatively high reflectance in the green band and low reflectance in the NIR band to separate water from soil and vegetated surfaces. This index can be regarded as the first index dedicated to water extraction. Even though the NDWI could suppress and remove non-water features to a large degree, Xu (2006) noticed that the index failed to suppress built-up land signals efficiently. Built-up area also has a relatively higher reflectance in the green band than in NIR, resulting in a positive NDWI value like water (Xu, 2006). Consequently, the extracted features could be a mixture of water and built-up land noises. Xu (2006) proposed a modified Normalized Difference Water Index (mNDWI), obtained by substituting the NIR band of NDWI with a SWIR band. The mNDWI results in negative values for built-up land and in positive values for water. Thus, the mNDWI can efficiently suppress and even remove the built-up noise which is not distinguishable in NDWI. Moreover, the mNDWI was proven to be more stable and reliable than NDWI, because the SWIR band is less sensitive to sediments or other active optical components than the NIR band. Because of its good performance, the mNDWI has become one of the most common WI used in open water delineation (Feyisa *et al.*, 2014; Ji *et al.*, 2009; Ogilvie *et al.*, 2015; Zhou *et al.*, 2017).

3.1.4. Thresholding Limitations

Thresholding is the most critical issue in using WI to extract water bodies. Since NDWI and mNDWI values for water are usually higher than zero due to the reflectance characteristics of water, a threshold of 0 is often applied to extract water from index images (Xu, 2006). However, Ji *et al.* (2009) suggested that adjustment of the threshold values could lead to higher extraction efficiency. Optimization of the threshold value is especially complex when studying the evolution of a water body across time, or when considering several water bodies from a given area because their reflectance characteristics vary a lot.

To deal with threshold optimization issue, researchers have developed a local adaptive thresholding scheme for existing indices (NDWI or mNDWI) (Li and Sheng, 2012) and new indices (AWEI and WI₂₀₁₅) which are less susceptible to spatio-temporal variation of reflectance (Feyisa *et al.*, 2014;

Fisher *et al.* ,2016). As an illustration, Feyisa *et al.* (2014) combined five bands from Landsat 5 TM to develop an Automated Water Extraction Index (AWEI) which is more resistant to shadow and built-up structures. One can distinguish two indices : one for images with no shadow (AWEIno shadow) and another for those with shadows from mountains, buildings and clouds (AWEIsh). A single optimal threshold of AWEI appeared to be capable of detecting water bodies from a time series Landsat imagery and optimal threshold was found to be less variable than the one of mNDWI.

3.1.5. Performance Comparison of Indices

Several empirical studies have examined the relative performance of indices through comparison of images classification results, collected by same or different sensors. Even though their findings are specific to the corresponding study, their results may give an indication about the most appropriate indices.

Jiang *et al.* (2014) found that the mNDWI and AWEIsh exhibited a larger accuracy than the NDWI and AWEInsh when used for the automatic extraction of three water bodies in China from Landsat Imagery with the AMERL methodology (Jiang *et al.*,2014). The results of Rokni *et al.* (2014) indicate the superiority of the NDWI in comparison with following others indexes (NDWI, mNDWI, AWEI, NDMI, WRI and NDVI) for the extraction of surface water of Lake Urmia in Iran from the Landsat data ETM+ 2000 and TM 2010 data. Zhou *et al.* (2017) compared open surface water bodies mapping approaches based on six water indices, (TCW, NDWI, mNDWI, sum of near infrared and two shortwave infrared bands (Sum457), AWEI, land surface water index (LSWI)) as well as three medium resolution sensors (Landsat 7 ETM+, Landsat 8 OLI, and Sentinel-2 MSI). Among the three sensors, Landsat 8 OLI and Sentinel-2 MSI performed better than Landsat 7 ETM+ for almost all of the algorithms, whereas the performance of Landsat 8 OLI and Sentinel-2 MSI were almost the same. Regarding the performance of the tested indices, the NDWI-based algorithms performed best followed by the algorithms based on AWEIsh, AWEInsh, mNDWI, SNN, mNDWI, LSWI, and TCW. These performances were assessed on a limited study area dominated by the pure open surface water body.

Even though the outcomes of each study are different, some indices often appear amongst the high-performance ones. This is the case of NDWI, mNDWI and AWEI. The latter would be probably be even more interesting for monitoring purpose since it has proven to be more stable across spatio-temporal variations of a given area.

3.1.6. Conclusion

Evolution in threshold-based methods enables the progressive reduction of the limitations regarding the water extraction accuracy, without completely overcoming them. In fact, combining two bands in ratios or indices improved the efficiency of water distinction in comparison with single band methods but only achieved a partial resolution of the spatio-temporal variability of the threshold. Attempts have been further conducted to develop local adaptive thresholding schemes for existing indices (NDWI or mNDWI) and new indices (AWEI and WI2015) which are less susceptible to spatio-temporal variation of reflectance. Considering the high number of available indices, some studies addressed the need for comparisons of indices performances. Their analysis revealed indices with promising accuracy for water detection : AWEI, mNDWI, NDWI. Even though the design and implementation of WI have been persistently improving (Jiang *et al.*, 2014), the lack of stability of the threshold

remains problematic, making it difficult to be used uniformly (Ji *et al.*, 2009). For this reason, the vast majority of the studies combine the threshold-based methods with other water extraction methods like classifiers-based methods, which are reviewed in the following section.

3.2. Classifiers-based Methods

Classifiers-based methods constitute the second broad category of techniques for water extraction. They can be used alone or in combination with threshold-based methods. In both cases, classifier-based methods enable to overcome some of the aforementioned limitations (Byoung *et al.*, 2015). The basic principle of image classification procedure is to automatically categorize all the units of satellite images into representative classes according to their resemblance (Lillesand *et al.*, 2015). Generally, classifiers-based methods can be divided into three catalogues : 1) based on the type of learning (supervised and unsupervised), 2) based on assumptions on data distribution (parametric and non-parametric) and 3) based on the spatial unit (pixel or object) (Jawak *et al.*, 2015).

3.2.1. Un-supervised and Supervised Classification Methods

Classifications may be differentiated in two categories (supervised or unsupervised) according to the learning approach of the algorithm handled by the analyst.

In supervised classification, the analyst supervises the pixel categorization process, by specifying to the computer algorithm, numerical descriptors of the land types present in a scene. To do this, representative sample sites of known cover types, called training areas, are used to compile a numerical “interpretation key” that describes the spectral (or others) attributes for each feature type of interest. Each pixel of the data is then compared numerically to each category in the interpretation key and labelled with the name of the category it resembles the most (Lillesand *et al.*, 2015). The largest source of error is the specification of numerical descriptors to the computer algorithm which have to be representative of all the within land-type diversity. The most popular classifiers falling under the category of supervised classification are Maximum Likelihood Classifier (MXL), Mahalanobis Distance (MhD), Support Vector Machines (SVMs), Spectral Angle Mapper (SAM), Parallelepiped (PP), Minimum Distance (MD), Decision Tree (DT), Random Forest (RF) and Nearest Neighbours (NN) (Jawak *et al.*, 2015).

In unsupervised procedures, the image data are first classified by aggregating them into clusters according to their similarities. The number of clusters has to be specified in advance to the algorithm by the analyst. Then, the land cover identity of each cluster is determined with the help of the analyst by comparing the classified image data to the ground reference data (Lillesand *et al.*, 2015). The limiting step is the post-classification in which the analyst has to merge or split the clusters and to label them using land surface information as a reference (Byoung *et al.*, 2015). Other common drawbacks of this procedure are the computational cost of the classification and the absence of a priori method to evaluate the ideal number of clusters (D’Andrimont *et al.*, 2017). The most popular classifiers that use this type of classification are K-means and Iterative Self-Organizing Data Analysis (ISODATA) (Jawak *et al.*, 2015). ISOSEG is another unsupervised algorithm widely used in relevant environmental remote sensing studies (Rizzi *et al.*, 2006 ; Quintano *et al.* 2005, 2006 ; Oliveira *et al.*, 2008 ; Shimabukuro *et al.*, 2009 ; Kawakubo *et al.*, 2009, 2011 ; Virdis, 2014).

Depending on the situation, one could prefer unsupervised procedures over supervised ones or the opposite. Generally speaking, supervised procedures tend to provide more accurate and meaningful

results in comparison with unsupervised techniques since their classes depict ground features (Paul *et al.*, 2017), whereas in unsupervised procedures, classes do not represent feature on ground and it is difficult to interpret them (Paul *et al.*, 2017). However, unsupervised procedures are usually better to cope with the lack of reference data (both ground truth and training samples) (PetitJean *et al.*, 2012b). Thus, in this particular case they are preferred over supervised procedures (Lu and Weng, 2007). This situation often occurs when dealing with change detection applications in the remote sensing domain (Jawak *et al.*, 2015; PetitJean *et al.*, 2012b) since reference data are not often available for each image captured.

3.2.2. Parametric or Non-parametric Classification Methods

Classifications may also be distinguished according to the presence of assumptions regarding the data distribution and the number of parameters required by the algorithm to characterize similarities among the data.

Parametric classifiers make a priori assumptions about the data distribution. These classifiers are characterized by an explicit underlying probability model, which provides a probability measure of any pixel in each class rather than simply a class value (Sharma *et al.*, 2013). This probability measure is based on statistical parameters such as mean and standard deviation. The number of parameters used to describe this similarity measure is finite and fixed by statistical laws. In contrast to parametric classifiers, non-parametric classifiers do not make any a priori assumption about the data distribution; distance is the only indicator of similarity amongst data. Moreover, the number of parameters serving for the characterization of distance measure is potentially infinite and tends to increase with the training set.

Parametric algorithms are MXL, PP, MD and ISODATA. Non-parametric algorithms are Artificial Neural Networks (ANN), Radial Based Function (RBF), SVM, DT and RF.

Parametric classifiers may perform well if the strong assumptions on the distribution of the data turn out to be correct and if the estimation of the model parameters is accurate (Caetano, 2017). Otherwise, classification accuracy can be strongly affected. Non-parametric methods can overcome these limitations since they are not constrained by assumptions regarding frequency distribution (Löw *et al.*, 2013). This is very useful because the data acquired from remotely sensed imagery usually have unknown distributions (Su and Huang, 2009).

3.2.3. Pixel-based or Object-based Classification Methods

Digital images classification can be distinguished based on the spatial units considered.

Traditionally classifications are performed using the pixel as spatial units. These classifications categorize all the pixels of an image into the thematic classes based on the spectral information of each pixel in the image. The spectral information used by each pixel is also termed as spectral signature, which is estimated by the relative reflectance in different wavelength bands. This pixel-based approach has been and is still being used for characterization and mapping the spatial extent of water bodies, among others (forests, urban, coastal, and wetland areas) (Hassan *et al.*, 2014). However, these pixel-based approaches present the following limitations and disadvantages : 1) low/limited efficiency of the classification, 2) the reliance on spectral information only, 3) sub-pixel problems, and 4) the speckle noise effect (Jawak *et al.*, 2015).

To overcome them, object-based image analysis (OBIA) approach has been thoroughly developed in the last two decades, by generating and analysing meaningful image objects instead of individual pixels (Yu *et al.*, 2006; Blaschke, 2010). The OBIA approach is conducted through a two-step process : (1) image segmentation by aggregating a number of individual pixels or image sub-objects to form larger objects (primitive objects) based on the homogeneity, intensity, and texture of each investigated image ; and (2) image classification and feature extraction (Jawak *et al.*, 2015). The processing unit of the classification is no longer a single pixel. The information used for the object classification is not restricted to spectral information but can be based on other features like form/shape, area/size, texture and content of the object (Kux and Souza, 2012). Ma *et al.* (2017) provided a comprehensive review on supervised object-based land-cover image classification. The result of the OBIA approach was successfully proved more accurate than that of the pixel-based approaches for land cover classification in recent studies (Dao and Liou, 2015) such as flood area delineation in the transboundary areas using the ENVISAT/ASAR and Landsat TM data (Mallinis *et al.*, 2011), and crop mapping using the multi-temporal Landsat imagery (Devadas *et al.*, 2012). Other applications of the object-based method for flood water and wetland mapping were introduced in (Dronovas *et al.*, 2015; Dronovas *et al.*, 2011; Frohn *et al.*, 2011). The object-based approach is more used with high spatial resolution images (<5m) but have also been proven to be effective in medium resolution imagery (<30m) (Dao and Liou, 2015; Bisquert *et al.*, 2010; Blaschke 2010).

3.2.4. Common Classifiers

According to Tehrany *et al.* (2019), comparative studies like Sharma (2013), Marjanovic *et al.* (2011), Tehrany *et al.* (2015), Althuwaynee *et al.* (2016) and Tien Bui *et al.* (2016c) have proven the superiority of machine learning (DT, ANN and SVM) over statistical techniques.

DT is the simplest non-parametric classifier algorithm. DT classifiers present a hierarchical configuration where, at each level, a test is employed on one or more attribute values. The outcome of this test may be a defined class or another decision node. The structure and criteria for the test at the decision node of the DT are both estimated from training data (Jawak *et al.*, 2015). The explicit and intuitive classification structure it provided constitutes one of its advantages, together with the effectiveness and robustness of the method when dealing with the noisy and non-linear relation among input feature and class labels (Friedl, 1997; Sharma *et al.*, 2013). DT have demonstrated significant performance in water bodies detection. As an illustration, Bertels *et al.* (2016) designed a water bodies detection algorithm (WBDA) based on DT classification. The WBDA proved to be effective for lake detection and monitoring using time series over Rift Valley in Ethiopia and is now used in the Copernicus Global Land Service to produce water map spatial resolution of 1 km (Proba-V) and 300 m (Sentinel-3). Bertels *et al.* (2016) also stated that the methodology would be compatible with higher resolution sensors as Sentinel-2. DT requires however large training samples and they tend to overfit them (Ghimire *et al.*, 2012).

ANN and SVM are more complex machine-learning classifiers. ANN are rules sets that simulate human learning processes through the establishment of linkages between the input and output data (Jawak *et al.*, 2015) whereas SVM aims to find a hyperplane that separates the dataset into a discrete predefined number of classes, consistent with the training examples. Both are designed to handle large volumes of data, including many bands and dates of observations. This capability along with their increased complexity in comparison with DT classification is resulting in a slightly increased

performance. Another advantage is the reduced dependence on training data, indeed they work well with small training datasets and are resistant to missing data (Mountrakis *et al.*, 2011). Regarding their drawbacks, they are computationally intensive and require the definition of input parameters (Jawak *et al.*, 2015). Another drawback is its “black box” character, meaning the explicit decision rules used to draw a given results remain unknown for the user (Rodríguez Galiano *et al.*, 2011). Finally, it is worthy to mention that the SVM tend to be more effective for binary outputs.

Machine learning classifiers can be combined to simultaneously take advantage of their strengths and further enhance a classification’s precision. The resultant algorithms are denoted as a ML ensemble classifiers.

Breiman’s Random Forest (RF) algorithm is an extremely popular ML ensemble classifier, which has recently started to gain prominence in water resources applications (Tyralis *et al.*, 2019). RF results from a bootstrap aggregation of classification or regression tree predictors (CART) with some additional degree of randomization. The bootstrap aggregation, also called bagging, refers to the combination of CART proposed in Breiman (2001) in order to alleviate their individual instability. The aggregation is done by taking the average of the predictions of the *ntree* trained models. The prediction is obtained by training *ntrees* times a CART using a new generated bootstrap sample each time. The randomization is performed to reduce the correlation between the trees and, consequently, reduce the variance of the predictions. Eventually, prediction in classification is performed by obtaining the majority class vote from the individual tree class votes. Increased predictive performance (2,80), fast computation, stable robust to noise have been demonstrated by RF.

RF also presents the advantage of allowing the estimation of variable importance, through the use of variable importance metrics which rank the predictor variables in terms of their relative significance. However, the provided information does not reflect the absolute performance of individual predictors in modeling the response variables. The two major variable importance metrics used in RF applications are : the mean decrease of the Gini index in classification and the mean decrease in accuracy. Another considerable asset of RF classifiers is the limited number of user-defined parameters. Only two parameters of RF algorithms need to be optimized : the number of trained trees *ntree* and the number of randomly selected predictor variables used at each node. For more information, Tyralis *et al.*(2019) list of appealing properties of RF presented in the recent literature. RF has largely proven to be effective for land cover classification, as well as for flood area mapping in urban environment (Feng *et al.*,2015).

3.2.5. Conclusion

To conclude, one could state that supervised procedure will be preferred over unsupervised ones for the meaningful classes they provide, except when there is a lack of reference data. Furthermore, the non-parametric methods should be adopted when the assumption of the data distribution is uncertain, even though they are more complex. The accuracy of parametric methods can be significantly reduced if the assumption turns out to be incorrect. Thank to the integration of additional information concerning the entities of interest (shape, area, texture, etc.), object-based classification has resulted in an improved accuracy for land cover classification but has also obviously added complexity in the processing of RS images. Consequently, pixel-based approach should be considered when spectral information is sufficient to reach the desired accuracy and if not, one should opt for object-based classification. Moreover, ML algorithms have demonstrated impressive performances when dealing

with few reference data and large amount of input data. Since running such complex algorithms is demanding in both computational power and time, ML classifiers are recommended only when large amount of data need to be processed with few reference data available, provided that resources of time and computational power are available. Among ML classifiers, the ensemble classifiers of RF is the most used.

3.3. Combination of Threshold and Classifier-based Methods

The combination of threshold and classifier-based methods consists in computing water indices without thresholding in order to enhance the distinction of water from background scene in the images to be classified. This significantly facilitates the classification. The following studies integrated this approach : Scarp (2016), Dao and Liou (2015), Mallinis *et al.* (2011).

4. Accuracy Assessment

The accuracy assessment is necessary to document the precision of satellite image classification results (Banko, 1998). It is usually conducted by selecting a sample of reference locations and comparing the classification at these reference locations to the ground truth. The reference sample should be selected independently of data used for training and/or developing the classification procedure. The most common means to express and visualize classification accuracy is the confusion matrix. This matrix lists the values for known cover types of the reference data in the columns and for the classified data in the rows (Banko, 1998).

Additionally, various metrics can be derived from the confusion matrix, also aiming at representing the classification accuracy (Stehman, 1997). The most basic descriptive metric is the overall accuracy (Equation 1).

$$OA(\%) = \frac{\sum_{i=1}^J n_{ii}}{\sum_{j=1}^J \sum_{i=1}^J n_{ij}} \quad (1)$$

where J = total number of classes and n_{ij} the number of observations predicted in the class j while belonging to the reference class i .

Besides the overall accuracy, the F1-score metric gives a comprehensive indication of the classification performance per class (Equation 4). This metrics is derived from the user's accuracy and producer's accuracy, also called precision and recall respectively. The producer accuracy (Equation 2) measures how well a certain area has been classified whereas the user accuracy (Equation 3) informs the user on how well the map represents what is really on the ground.

$$\text{Precision}(\%) = \frac{n_{ii}}{\sum_{j=1}^J n_{ij}} \quad (2)$$

$$\text{Recall}(\%) = \frac{n_{ij}}{\sum_{i=1}^J n_{ij}} \quad (3)$$

$$\text{F1-score}(\%) = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

Although informative, measures implying the percentage of cases correctly classified have often been

criticized. (Foody, 2002). A major problem for some users is that some cases may have been allocated to the correct class purely by chance (Congalton, 1991 ; Pontius, 2000 ; Rosenfield Fitzpatrick-Lins, 1986 ; Turk, 1979). To accommodate for the effects of chance agreement, Cohen's Kappa coefficient is often been used. The Kappa coefficient is a measure of the difference between the actual agreement between reference data and an automated classifier and the change agreement between the reference data and a random classifier (Lillesand *et al.*, 2015).

However, the Kappa coefficient has been the subject of controversy. Some commentators argue that it should be adopted as a standard measure of classification accuracy (Smits *et al.*, 1999) whereas, others like Pontius and Millones (2011) recommended that the profession abandons the use of Kappa indices for purposes of accuracy assessment and map comparison. Instead, Pontius and Millones (2011) suggested two other summary parameters : quantity disagreement Q and allocation disagreement A for category-level accuracy assessment (Warrens, 2015).

$$Q = \frac{1}{2} \sum_{i=1}^J (p_{i+} - p_{+i}) \quad (5)$$

$$A = \left[\sum_{i=1}^J \min(p_{i+}, p_{+i}) - \sum_{i=1}^J p_{ii} \right] \quad (6)$$

The most appropriate parameter to report the classification result, among the bewildering choice of parameters, depends on the pursued objective. However, because it is difficult to anticipate objectives and accuracy needs of all eventual users of a map product and because no consensus has been reached on which measures are appropriate for a given objective of accuracy assessment among scientists, the best course of action is to report the full error matrix along with the sampling design used to collect the reference data (Stehman, 1997). In addition, Foody (2002) states that at least one quantitative metric of classification accuracy together with appropriate confidence limits need to be provided in addition to some general level of accuracy measure.

5. Water Monitoring

Detecting and monitoring water surface change over regional areas are usually achieved by analyzing a single time series of satellite images (STIS) through the process commonly called 'change detection' (Al-doski *et al.*, 2013 ; Coppin and Bauer, 1996). Change detection is defined as the process of identifying differences in the state of an object or phenomenon by observing it in different times (Singh, 1989).

Change detection methods can be distinguished into two approaches depending on the chronological order of the classification and the comparison phase. The first one consists in classifying the images of the time series and then compares the classification results, whereas the second approach directly compares the images of the time series and then, classifies the comparison results into different status of changes (Huang *et al.*, 2018). These approaches are referred to as post- and pre-classification comparison method respectively. The classifier used for change labelling are not specific to change detection (Coppin *et al.*, 2004). Since these classification techniques already been reviewed in Section 3, the following section focused on the comparison methods.

Post-classification comparison involves independently produced spectral classification results from each end of the time interval of interest, followed by a pixel-by-pixel or segment-by-segment comparison to detect changes in cover type. By properly coding the classification results, a complete matrix of change is obtained, and change classes can be defined by the analyst (Coppin *et al.*, 2002). The fact that images are classified independently can be viewed as an advantage because it minimizes the problem of radiometric calibration between dates. The accuracy of the post-classification comparison is totally dependent on the accuracy of the initial classifications. Surface water change detection is usually conducted according to post-classification techniques (Rokni *et al.*, 2014). In fact, numerous studies used this technique to estimate river basin area change (Du *et al.*, 2012), lake volume change detection (Lopez-Caloca *et al.*, 2008), wetland water quality and area change (Zhou *et al.*, 2011; Xu *et al.*, 2010) and coastline change detection (Alesheikh *et al.*, 2007). This approach, however, is usually very demanding in terms of classification work and errors may accumulate when comparing the results of classification (Huang *et al.*, 2018). A growing number of studies have therefore been exploring methods belonging to the second category.

The second category comprises methods like image differencing (Wang, 2004; Wang *et al.*, 2002), image rationing, image regression, principal component analysis (Gianinetto *et al.*, 2006; Rokni *et al.*, 2014) change vector analysis (Huang *et al.*, 2016; Landmann *et al.*, 2013) and other linear data transformation.

Initially devoted to bi-temporal analysis, images differencing/rationing/combination and regression were extended to multi-temporal analysis (Cohen *et al.*, 1998). However, using these methods to compare two or several images of a time series appears more as an ad hoc technique than an effective method to handle whole time series. Furthermore, linear data transformation methods present the advantages of the being simple to use and exploiting the whole data but are also unable to depict temporal behaviours and are intolerant to irregular sampling. For these reasons, all the aforementioned methods are not able to make a consistent temporal analysis of an entire satellite images time series (SITS). When the study has to consider the whole data and the time order in time series, researchers generally turn to trends analysis methods.

Trends analysis methods analyse entire SITS in a single step exploiting the whole-time sequence information and the time ordering (Débonnaire *et al.*, 2017). Maintaining the temporal order enables the mining of structured information and the analysis of evolution behaviours from the image series.

This can be achieved by extracting frequent sub-sequences of radiometric evolutions with methods like frequent pattern mining (Petitjean *et al.*, 2010; Mougél *et al.*, 2013; Julea *et al.*, 2011; Romani *et al.*, 2013; Xue *et al.*, 2014) and frequency analysis (Débonnaire, 2017, Andres *et al.*, 1994; Celik, 2009 and Celik and Ma, 2011). Frequency analysis and frequent pattern mining extract numerous meaningful patterns which are difficult to filter in order to understand changes that occurred over an area (Petitjean, 2012a). While able of handling long time series and are robust to noise, they are unable to handle irregularly sensed data over the image time series. This is problematic since the regular acquisition remotely sensed data is often compromised operational or meteorological factors.

More simple approaches rely on graphs to represent temporal evolution of pixels or objects across the time series and are performing very well (Güttler *et al.*, 2017; Güttler *et al.*, 2014; Rejichi and Chaabane, 2012).

6. Conclusion

In conclusion, this literature review highlighted, through an overview of the available studies (Section 1), that water presence was the main common element to be looked at when remote sensing is used for detection and monitoring of object similar to SB. This further guided the objective of this master thesis.

Regarding methodological choices, medium resolution sensors are the one providing sufficiently large spatial coverage and high frequency of observations to monitor SB at the regional level. Sentinel-2 is the most convenient as it provides cost-free images with enhanced temporal and spatial resolution and coverage in comparison with Landsat, SPOT and ASTER. When looking among the method to detect water, a promising option would be to perform an enhancement of the Sentinel-2 images with water-related indices, followed by a classification. Since it is foreseen that spectral information suffice from monitoring water, there is no need to opt for object-based classification. A preference goes for the RF classifier because it has already demonstrated impressive performances. Finally, the monitoring of water could be studied and represented by means of graphs representing the temporal variation of classified pixels.

Chapter III. Objectives

The overall objective of this research is to evaluate the feasibility of monitoring the hydrological functioning of SB using Sentinel-2 time series, while focussing the research on the province of Walloon Brabant as a study case.

To meet this goal, it appears, quite early in this study, necessary to foresee a step-by-step approach relying on three more specific and sequential objectives.

The first objective to be pursued is to prepare the SB dataset and characterize ex-ante the detectability of each SB, based on the adequacy between the SB's characteristics and the Sentinel-2 images limitations (spatial and temporal resolutions). Performing this characterization on the SB should enable to select the readily observable SB as a suitable base to conduct the attempt of monitoring. Furthermore, this should allow the identification of transposable parameters to determine ex-ante the detectability of the SB located in other geographical scopes the one chosen in this study.

The rest of this research investigates more specifically the potential of Sentinel-2 time series in monitoring the temporal variation of the water surface of observable SB, through two further objectives.

The second objective aims to calibrate and validate a classification model which is able to exploit the information provided by three Sentinel-2 calibration/validation images, in order to detect the presence of water inside the readily observable SB at the corresponding times.

Finally, the third objective is to investigate whether the information provided by the Sentinel-2 time series enable the developed water classification model to depict the temporal variation of the water surface of observable SB. This further related to assessing if the Sentinel-2 time series provides the required information to derive the operational state of observable SB.

Chapter IV. Study Area

The geographical scope of this study is the province of Walloon Brabant. This specific choice has, among other factors, been dictated by the fact that this region was also the scope of the initial similar project conducted in partnership by the Walloon Brabant Province and the UCLouvain - Earth Life Institute which by the way, resulted in the best dataset of SB available to date.

The province of Brabant Walloon is located in the center part of Belgium [Lat. 50,48–50,72N, Long. 4,08–5,16 E] and covers about 1097 km² ($\approx 3,6\%$ of Belgium) (Figure 3).

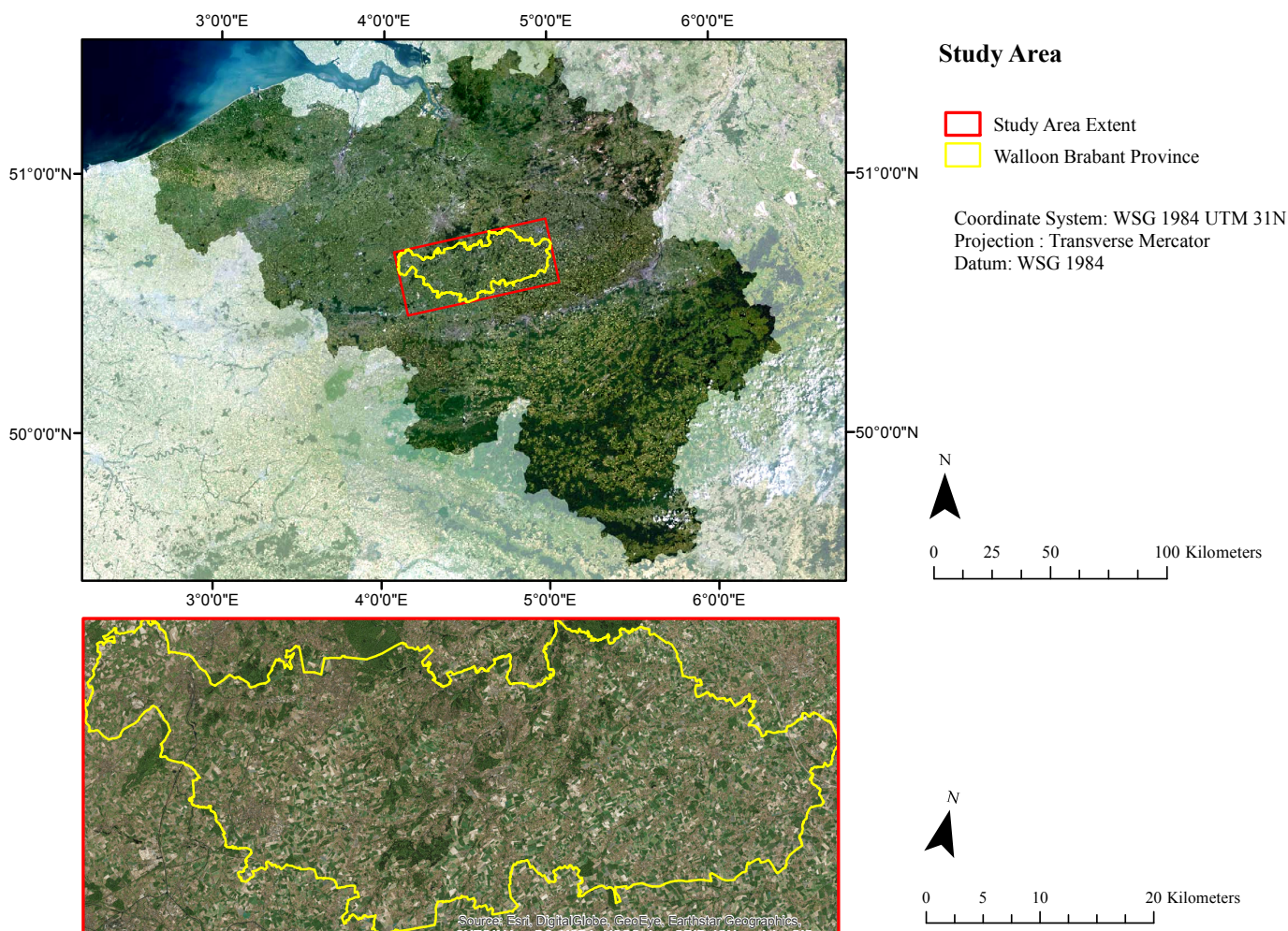
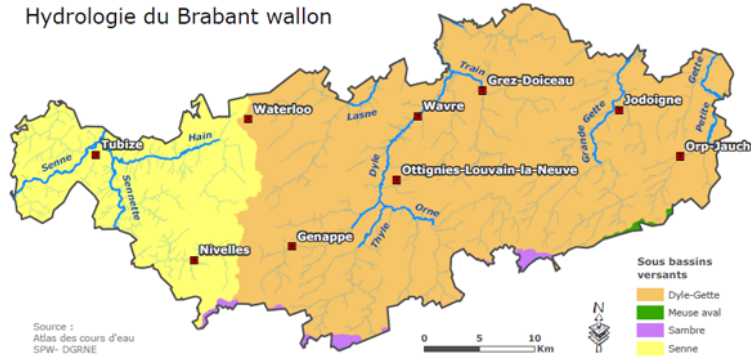
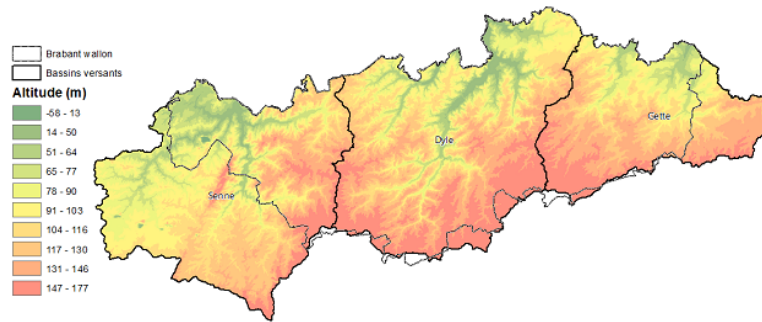


FIGURE 3 – Study area extent

Its territory is crossed by five sub-watersheds (Figure 4.a). The Seine is accounting for 23% of the territory (West) whereas both Dyle and Gette sub-watersheds are representing 76% of the territory (East). The remaining percent falls in the Sambre and Meuse Aval sub-watersheds. The subsequent study focuses on the Seine, Dyle and Gette sub-watersheds because they offer a manageable size to analyse and above all, in contradiction with the other sub-watersheds, they benefit from an available dataset of recorded SB. The three sub-watersheds cover a total area of 1521 km² and display similar elevation profiles with minimum varying between 30 and 50 m in the alluvial plains and maximum on crests and split lines rising up to 140-180 m (Figure 4.b) (Walloon Brabant Province, 2019a).



(a) Sub-watersheds and principal watercourses



(b) Elevation of sub-watersheds

FIGURE 4 – Sub-watersheds and principal watercourses of Walloon Brabant Province and their corresponding elevation (Source : Wallonie Public Service (SPW) and Wallonia’s Operational Directorate General for Agriculture, Natural Resources and the Environment)

According to the national meteorological data (IRM, 2019), the Walloon Brabant has a temperate oceanic climate. In fact, the study zone is influenced by the Gulf Stream and is subject to ocean disturbance while at the same time benefiting from its thermal inertia. This is resulting in moderate temperatures, high cloudiness and frequent but sparse rainfall. It rains on between 130 and 145 days (>1 mm/day) per year in Walloon Brabant, for a mean annual precipitation ranging from 700 mm to 1000 mm, resulting in a relatively evenly distributed precipitation in time. The total precipitation is representing a water intake of about 0,7-1 billion cubic meters per year. During the summer months (June to September), temperatures are increasing (higher than 15°C) while decreasing (fewer than 15°C) during the winter period (December to March).

The Walloon Brabant Province (2019a) reported having recurrently experienced floods, also defined as the abnormal water coverage of lands (Directive 2007/60/CE, Code de l’Eau). In most instances, these are isolated local events which are not relayed in the media. However, this study zone has also been hit by large-scale weather events like heavy rainfall or storms, causing floods. Such harmful events may be recognized as public calamities by the Federal Government (WBP, 2019a). Their occurrence frequency is directly related to periods of heavy rain in the province. Among the 36 flood events acknowledged as public calamities in Walloon Brabant between 1995 and 2018, almost half of them occurred during summer months (1 during June, 9 during July and 7 during August) (WBP, 2019a; SPW, 2019; Federal Public Service Justice, 2019). November, December and January are also prone to flooding (9 events). From a historical perspective, the flood events are not evenly distributed across time. It seems that flood disasters have worsened and became more frequent over time. Rea-

sons therefore may be multifold : increased frequency of intense precipitations, increased urbanization leading to soil sealing, etc. (WBP, 2019).

In order to mitigate flood risk and damages, the province has built the first SBs in the years 1970 as a low-cost solution to store the running off stormwater. The construction of SB has continued up to now.

The management¹ of those SB is fragmented and executed by different authorities at different levels. One can distinguish three broad types of SB managers depending on their location.

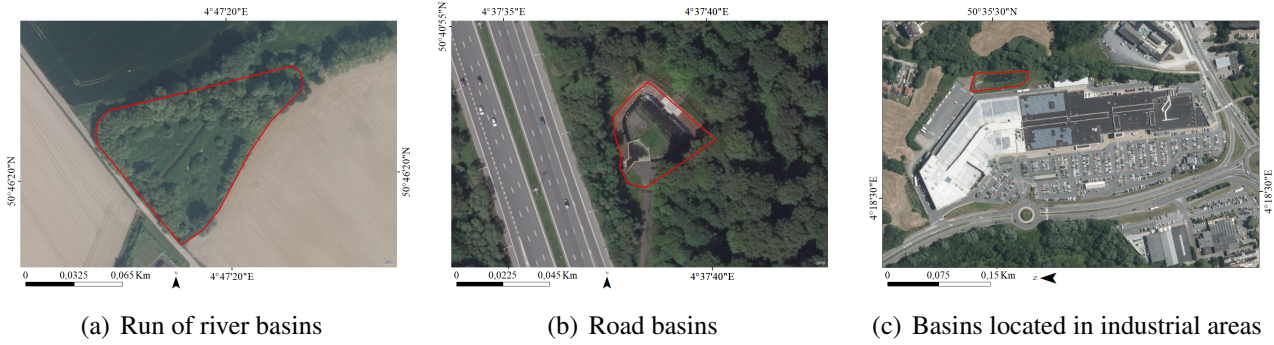


FIGURE 5 – Illustrations of the three common locations of SB : beside or run of river roads as well as next to urban area (Roads or other urban area)

Run of river SB or SB associated with watercourses usually fall under the responsibility of the administrators of the water system (Figure 5.a). In Belgium, non-navigable streams are divided into three categories according to the size of the watershed that feeds the concerned streams. Each category is managed by a separate and different entity (Figure 6). Consequently, the SB manager depends on the category of the watercourse it is connected to.

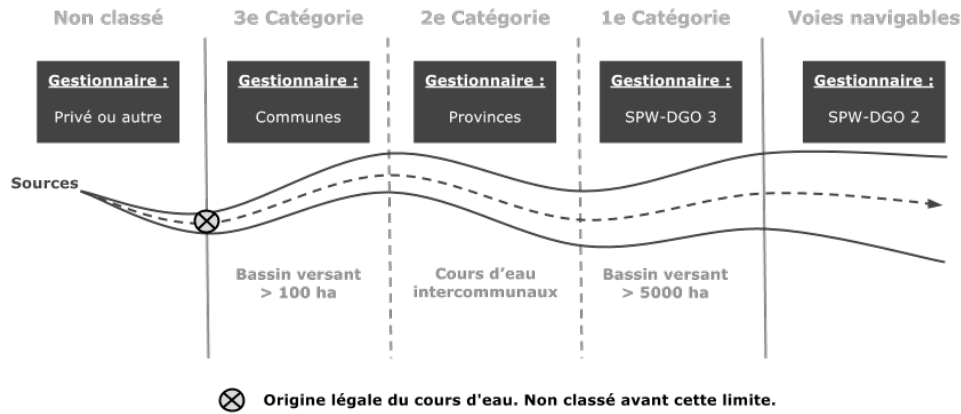


FIGURE 6 – Categorization and corresponding administrators of the water system in Belgium (Walloon Brabant Province, 2019)

Besides stream related SB, numerous SB are located along the roads (Figure 5.b). By virtue of their position, their management is ensured by the administrators of the road network. In Belgium, the management of the road network is decentralized. According to the Public Service of Wallonia for Mobility and Infrastructure (2019), the road network consists of two groups of roads : 1) regional roads managed by the department of Wallonia's Operational Directorate General for Roads and Bul-

1. For the purpose of this study, the general notion of management covers as well governance and specific management.

dings (DGO 1) and partially by the SOFICO ; 2) the provincial and communal roads administrated by the relevant provinces and communes.

Finally, SB can also be managed by private operators. This is frequent when SB were created to compensate the soil sealing caused by the construction of urban areas such as building complex or industrial areas (Figure 5.c).

This demonstrates the high number of different potential managers of SB over the province of Walloon Brabant and the difficulty to determine, a posteriori and with no additional information, the effective manager. This could further justify the actual difficulties to settle an integrated management of SB and the problematic situations stemming from the inability to determine the manager of basins without the help of the archives.

Furthermore, SB management obligations should logically have to be described in the legislation dealing with flood risk management. The latter is covered by the framework PLUIES, the flooding European Directive (2007/60/CE) and the Water Code. However, it can be rapidly noted that these documents remain very vague on those aspects and that very few clear legal obligations are identified.

According to Mr. Libert, Director of hydraulic research at SPW (interviewed by telephone on October 1, 2018), the competent administration is accountable for the construction and the maintenance of SB. In that sense, the competent administration prepares the contract specifications for the construction and maintenance work and launches the public bid. Further on, the administration will be in charge of the supervision and validation of the work realized by the successful bidder. The frequency and the priority of maintenance work are left to the manager's discretion. This absence of strict obligation might be one reason for the current maintenance backlog of SB managed by the SPW.

To conclude, beside the fact that it is scope of the initial similar project of SB data centralization, the province of Walloon Brabant constitutes a relevant study case because of its vulnerability to flooding as well as its relatively high level of financial investment in remediation measures, leading to the construction of man-made SB. This geographical scope is also appropriate because of the existing challenge to settle an integrated management of SB at the provincial scale. In fact, the province encounters difficulties to coordinate the SB management between the multiple SB managers, especially since the SB manager are not subject to strict obligations about the SB management.

Chapter V. Material

This chapter lists, from a material perspective, the ground-collected and remote sensing data sources and the softwares used in the realization of this thesis.

1. Ground-Collected Data

1.1. SB Reference Dataset

Ground-collected data originating from provincial platform of flood risk management project were firstly used to locate the SB. This provincial platform of flood risk management project results from a collaborative effort between the province of Walloon Brabant and UCLouvain - Earth and Life institute initiated through a management contract in 2012-2014 and 2015-2017 (Walloon Brabant Province, 2019). This project aimed to define a comprehensive strategy to improve flood risk management and has led to the establishment of a platform of knowledge and expertise. Since 2017, the development and update of the platform constitute a core responsibility of the province.

This platform is of particular interest for this thesis since it includes a geospatial SB database. To be exact, other flood mitigation infrastructure (temporary immersion area, ridges/valley and underground SB) are recorded into the platform. However, only the open and artificial stormwater basins² were considered for the purpose of this study. Their inventory was undertaken with the aid of questionnaires and interviews with representatives responsible for the upkeep of SB. As a result, this cartographic database currently constitutes the best reference dataset available, even though it was acknowledged to be incomplete and non-validated in the final project report (Javaux *et al.*, 2014). The lack of completeness of this dataset appeared to be caused by the lack of cooperation of some communes, by the absence of records in the archives or even by the loss of the archives.

This geospatial database can be consulted in the form of an interactive map at the following address : <http://www.brabantwallon.be/inondations/>. Access to the non-public part of the platform and an exploitable text file were provided by the province following an interview with Alexandra Bauwens (Agent in charge of the control of floods) and Vincent Balthazar (Director of the mapping and hydrology department of WBP). The version of the dataset used in the following study dates from November 14, 2018 (day of the interview) and report 156 SB out of 249 flood mitigation infrastructure identified in Walloon Brabant. The provided text file stores the following information for each of the 156 SB :

- SB registration (submission day, user ID)
- SB identifier (SB ID, name of SB)
- SB location (latitude (y_{lat}), longitude (x_{long}), address, commune)
- SB design (creation date, type of facility, type of evacuation/discharge, presence of a gate, presence of a spillway, name and category of the associated watercourse, SB volume, SB manager ID and operating state)

2. For the sake of simplification, the so-called "Open and artificial stormwater basins" in the platform are referred to as SB in the rest of this thesis.

1.2. Rainfall Time Series

Rainfall Time Series were used to support the interpretation of the reactivity of SB to the rainfall occurred in a neighbouring zone. These data are originating from the network of agrometeorological stations called Pameseb. The latter results from an initiative launched in 1989 by the UCLouvain laboratory in rural engineering. The province of Walloon Brabant comprises three stations located in Louvain-la-Neuve, Baisy-Thy and Roux-Miroir. However, only the data from the station of Louvain-la-Neuve (located at [Lat. 50.668962 N, Long. 4.642065 E]) were accessible free of charge.

1.3. Time Series of Water Levels Inside SB

Measurements of water levels inside SB were used as independent data to validate the obtained temporal evolutions of water. The latter were collected by the water-level network called Aqualim and managed by the department of Wallonia's Operational Directorate of non-navigable streams. Nicolas Feltz provided access to the non-public part of the Aqualim platform (<http://aqualim.environnement.wallonie.be/login.do?time=1565597454893>) through which the evolution of water levels have been obtained for 13 SB of the province of Walloon Brabant.

2. Remote Sensing Data

2.1. Orthophotoplans

The first type of remote sensing data were provided by the Walloon Public Service (SPW). The SPW takes periodically aerial photographs of the Walloon Region territory using a vertically positioned aerial-born sensor (at 3000-4000 metres in altitude). These aerial images are orthorectified and mosaicked in order to produce orthophotoplans. The resulting orthophotoplans cover the Walloon territory with a spatial resolution of 25 cm. Four spectral bands are available (green, red, blue and near infra-red band). The cartographic projection of these orthophotoplans is the Belge Lambert 1972. Collections of orthophotoplans exist for the following years : 1971, 1994-2000, 2006-2007, 2009-2010, 2012-2013, 2015, 2016, 2017 (partial coverage) and 2018. The latter are freely accessible online on the WalOnMap site (Geoportail of Wallonia, 2018). The orthophotoplans were used when a higher level of detail than Sentinel-2 was required. This comprises the delineation of SB and the assessment of the visibility of their surface.

2.2. Sentinel-2 Image Times Series

2.2.1. Description of Sentinel-2 Images

The second type of remote sensing data were provided by the Copernicus Sentinel-2 mission conducted by the European Space Agency (ESA). This source of remote sensing data was chosen for its long-term continuity of the current SPOT and Landsat missions, as well as the free and open access images it provides. More specifically, this mission comprises a constellation of two satellites, Sentinel-2A and Sentinel-2B, placed in the same sun-synchronous orbit and phased at 180° to each other. These satellites were launched by the ESA in June 2015 and March 2017 respectively. Their concurrent work enables the reduction of the revisit time from 10 days at the equator with one satellite to 5 days with 2 satellites under cloud-free conditions (which results in 2-3 days at mid-latitudes). Their wide swath width (290 km) and their high spatial resolution (10 m) provide the unprecedented possibility to work at both high precision and regional spatial scale. The collected images are divided

into tiles projected and referenced in the UTM cartographic projection and the WGS84 coordinates system respectively (ESA, 2019a.).

The Sentinel-2 MultiSpectral Instrument senses signals reflected from the earth in order to provide reflectance measurement in 13 spectral bands covering the visible and IR spectral range, with 10 m, 20 m and 60 m pixels (Table 2) (ESA, 2019b). In this study, the three 60 m bands were not used due to their coarse spatial resolution and because of their specific usage for atmospheric characterization and correction. The remaining ten bands were selected for the subsequent analysis.

TABLE 2 – Spectral and spatial resolution of the Sentinel-2 MSI bands (* denotes bands not considered in study).

Band	Central Wavelength (μm)	Bandwidth (μm)	Spatial Resolution (m)
Band 1 – Coastal aerosol*	442	21	60
Band 2 – Blue	492	66	10
Band 3 – Green	559	36	10
Band 4 – Red	664	31	10
Band 5 – Vegetation red edge	704	16	20
Band 6 – Vegetation red edge	740	15	20
Band 7 – Vegetation red edge	780	20	20
Band 8 – NIR Wide	832	106	10
Band 8A – NIR Narrow	864	22	20
Band 9 – Water vapour*	945	21	60
Band 10 – SWIR – Cirrus*	1376	30	60
Band 11 – SWIR 1	1619	94	20
Band 12 – SWIR 2	2200	185	20

2.2.2. Image Acquisition

According to the aforementioned elements, the Sentinel-2 satellites seem appropriate to support monitoring of SB surface changes. Therefore, Sentinel-2 Level-1C (top of the atmosphere) images were downloaded for the two tiles covering the study area and over the years 2016, 2017 and 2018. Their download has been automated through the Sentinel-2 for Agriculture (*Sen2Agri*) software. The resulting Level-1C time series for the tile UES was composed of 366 images spread between October 1, 2016 and December 12, 2018. For the UFS tile, the resulting Level-1C time series was composed of 238 images spread between October 5, 2016 and December 9, 2018. The time series comprises two co-existing temporal resolutions (10 and 5 days) since it benefits from the reduction of revisit time made possible by the launch of Sentinel-2B in July 2017. The time series counts significantly more UES tiles because the Sentinel-2 satellites frequently acquire only a part of this tile due to orbit coverage at high latitude. This is resulting in an increased number of partially covered tiles containing missing values at the non-acquired locations (Figure 7).

2.2.3. Atmospheric Correction

The pre-processing of Level-1C images was conducted using the *Sen2-Agri*'s Level-2A processor, based on the Multi-sensor Atmospheric Correction and Cloud Screening (MACCS) chain. The latter performs atmospheric corrections and produces bottom-of-atmosphere reflectance values with snow, large water bodies, cloud and cloud shadow masks (*Sen2Agri* and ESA, 2018). Since the pre-processing was not applied when the cloud coverage is higher than 90%, the number of informative

images in the Level-2A time series is reduced in comparison with the Level-1C one. Over the UES tile, the Level-2A time series comprises 155 images between October 1, 2016 and October 31, 2018. Over the UFS tile, the Level-2A time series comprises 100 images for the UFS between October 5, 2016 and October 28, 2018.

2.2.4. Clouds and Missing Values Removal

Each Level-2A Sentinel-2 image was further pre-processed to remove clouds or missing values resulting from the absence of valid reflectance (Figure 7) in the initial Sentinel-2 surface reflectance image.

To this end, two masks were developed to identify the position of clouds and missing values respectively in the initial Sentinel-2 image. The cloud mask was generated based on a conditional selection applied on the cloud mask provided by the Sentinel-2 for Agriculture (*Sen2Agri*) system through the atmospheric correction process. The cloud mask provided by *Sen2Agri* contained various levels of cloud (and their corresponding shadow) intensity. For the ease of calculation, these levels were simplified to binary values (1-0) indicating the presence of clouds and their corresponding shadow. Through this simplification, it has been assumed that presence of clouds or their corresponding shadow would significantly affect reflectance irrespectively of their intensity. The missing value mask was built following a conditional selection (different from -10000) applied on pixel values, belonging to one spectral band of the 10 m resolution bands. This approach was adopted because the raster on which the conditional selection was applied provided the required information for the creation of masks compatible with all the raster bands.

Since the two masks were characterized by a spatial resolution of 10 m, they were successively and directly applied on the four spectral bands with 10 m spatial resolution. The spectral bands with 20 m spatial resolution had to be downsampled and resampled to 10 m prior to the successive application of the two masks. The resampling was done according to the bilinear interpolation operator.

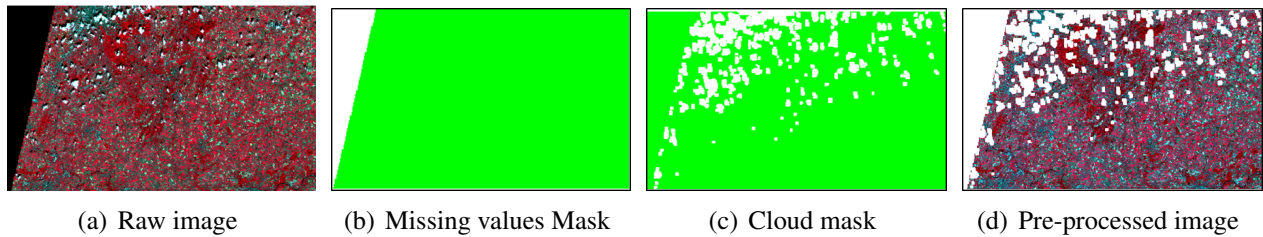


FIGURE 7 – Workflow of clouds and missing values removal

After the successive application of the two masks on each Sentinel-2 image of the time series, all the bands of the Sentinel-2 time series were characterized by a common value (NA) for pixels subject to the presence of clouds or missing values resulting from the absence of valid reflectance in the initial Sentinel-2 surface reflectance image (Figure 7). This newly assigned common value indicates the impossibility to use the reflectance measurements from the initial Sentinel-2 image.

2.2.5. Water Indices Calculation

The 10 pre-processed spectral bands further served the purpose of water indices calculation. Literature review enabled the selection of the water indices presented in Table 3. Since the vast majority of indices were developed using the spectral bands of other missions than the Sentinel-2 mission, the

Sentinel-2 spectral bands used to calculate the indices were selected to be the most consistent with their original definition.

The 10 spectral bands of Sentinel-2 images resampled at 10 m resolution and the 11 water spectral indices listed here after constitute the set of explanatory variables given to the classification model to enable the detection of water. These explanatory variables are also referred to as "feature" in the rest of this document.

TABLE 3 – Definition of the selected water indices

Index	Equation
Normalized Difference Water Index (Gao, 1996)	$NDWI = \frac{(GREEN - NIR)}{(GREEN + NIR)}$
modified Normalized Difference Water Index (McFeeters, 1996)	$mNDWI = \frac{(GREEN - SWIR_{1,65})}{(GREEN + SWIR_{1,65})}$
Automated Water Extraction Index _{shadow and dark surfaces} (Feyisa <i>et al.</i> , 2014)	$AWEInsh = 4(GREEN - SWIR_{1,65}) - (0,25 \times NIR + 2,75 \times SWIR_{2,13})$
Automated Water Extraction Index _{no shadow and dark surfaces} (Feyisa <i>et al.</i> , 2014)	$AWEInsh = BLUE + 2,5 \times GREEN - 1,5(NIR + SWIR_{1,65}) - 0,25 \times SWIR_{2,13}$
Land Surface Water Index (Chandrasekar <i>et al.</i> , 2010)	$LSWI = \frac{(NIR - SWIR_{2,13})}{(NIR + SWIR_{2,13})}$
modified Land Surface Water Index (Kwak, 2014)	$mLSWI = \frac{(1 - NIR - SWIR_{2,13})}{(1 - NIR + SWIR_{2,13})}$
Moisture Stress Index (Hunt <i>et al.</i> , 1989)	$MSI = \frac{SWIR_{1,65}}{NIR}$
Sentinel Water Mask (Milczarek <i>et al.</i> , 2017)	$SWM = \frac{(BLUE - GREEN)}{(NIR + SWIR_{2,2})}$
Value (Pekel, 2014)	$V = \max(SWIR_{2,2}, NIR, RED)$
Saturation (Pekel, 2014)	$S = \frac{(V - \min(SWIR_{2,2}, NIR, RED))}{V}$
Hue (Pekel, 2014)	$H = \begin{cases} 0 & \text{if } V = \min(SWIR_{2,2}, NIR, RED) \\ (60^\circ \times \left(\frac{NIR - RED}{V - \min(SWIR_{2,2}, NIR, RED)} + 360^\circ \right) + \text{mod}(360^\circ)) & \text{if } V = SWIR_{2,2} \\ (60^\circ \times \left(\frac{RED - SWIR_{2,2}}{V - \min(SWIR_{2,2}, NIR, RED)} + 360^\circ \right) + \text{mod}(360^\circ)) & \text{if } V = NIR \\ (60^\circ \times \left(\frac{SWIR_{2,2} - NIR}{V - \min(SWIR_{2,2}, NIR, RED)} + 360^\circ \right) + \text{mod}(360^\circ)) & \text{if } V = RED \end{cases}$

2.2.6. Creation of Calibration and Validation Images

Three images were selected from the Sentinel-2 time series to serve the purpose of calibration and validation of the classification model. Their selection has been dictated by the fact that their UES and UFS tiles were providing a cloud-free coverage of the study area at the time of the acquisition. The three selected Sentinel-2 images were taken at the following dates : July 15, 2018; September 18, 2018 and October 10, 2018. In virtue of their function, these images are referred to as "calibration/validation" images in the rest of this work.

These images were initially divided into two tiles, as the time series from which they have been extracted. However, it was crucial to calibrate/validate the model on images that cover the entire study area. Therefore, the UES and UFS tiles corresponding to each acquisition date were merged in order to produce three Sentinel-2 images bigger than study area. Their extent was further reduced to

the minimal bounding rectangle encompassing the study area (illustrated in red in Figure 3), in order to limit the computation time on raster operations.

Such merger was not performed on the whole time series because the UES and UFS tiles were not always collected at a same time. Consequently, their merging into one image per date would have resulted in a mix of single tile and two tiles images over the time series. This was not desirable for the process's automation. The monitoring of the water inside SB was thus performed considering the time series of one tile at a time.

3. Softwares

The *ArcGIS* software has been used for data visualisation, polygon delineation, cartographic projection and map production. The license provided by UCLouvain gave access to the following softwares : Desktop Standard, Network Analyst, 3D analyst, Spatial Analyst, Geostatistical Analyst and Tracking Analyst.

In addition, the *Sen2Agri* open-source software, developed by the UCLouvain laboratory in Geomatics and funded by ESA, has been used to automatically download the Level-1C Sentinel-2 image time series and pre-processed it into Level-2A Sentinel-2 image time series.

Finally, the *R Studio* software working with the programming language Markdown has been used to achieve automation of the operations described in the following sections. The following packages were used :

- For raster processing : 'raster', 'rgdal', 'rgeos',
- For classification : 'randomForest', 'caret',
- For features selection : 'VarSel',
- For statistics computation : 'stats',
- For visualisation/charting : 'ggplot2'.

Chapter VI. Methods

This chapter presents the methodology settled to assess the feasibility of monitoring the water inside observable SB using Sentinel-2 time series. It starts with an explanation of the overall approach (Section 1) and continues with a comprehensive description of the steps implemented to successively meet the three pursued specific objectives (Sections 2, 3 and 4).

1. Overall Approach

The overall approach can be conceptually divided into three steps, corresponding to the three pursued specific objectives (described in Chapter III).

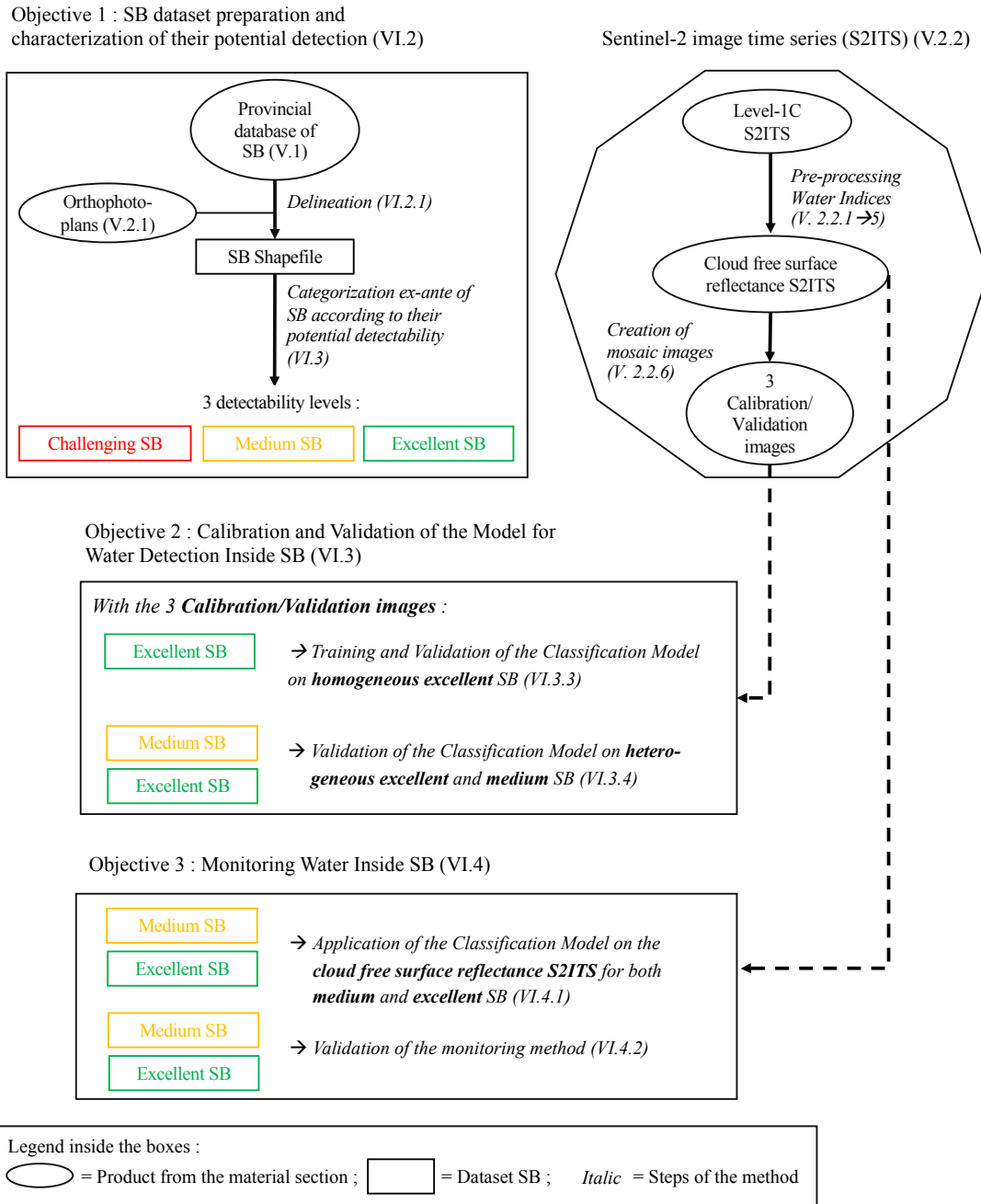


FIGURE 8 – Workflow of the overall methodology established to monitor the temporal variation of water presence inside observable SB

The first step of this research was firstly dedicated to the preparation of the SB dataset provided by the province of Walloon Brabant. This preparation was performed through a delineation of the SB based on their spatial coordinates recorded in the provincial database and the orthophotoplans, in order to create a SB shapefile. The delineation highlighted that some SB were barely observable by means of orthophotoplans, so it became very likely that Sentinel-2 images (more constraining in terms of spatial resolution than orthophotoplans) would not permit the detection of all the registered SB. Consequently, the second part of this step has been devoted to the ex-ante characterization of the detectability of each SB, based on the adequacy between the SB's characteristics and the Sentinel-2 images' limitations. This process enabled the categorization of the SB into 3 different levels of potential detectability : challenging, medium and excellent. The latter served the selection of the readily observable (medium and excellent) SB as a suitable base to conduct the attempt of monitoring. The challenging SB were deemed too constraining for attempting the application of the monitoring via the method developed hereafter.

Before attempting the strict monitoring of observable SB, a second step was taken towards the development of a classification model, which is able to detect the presence of water in those SB. A choice was made to calibrate and validate this model on three calibration/validation Sentinel-2 images. The training and validation of the classification model was firstly conducted on homogeneous excellent SB, since the water detection is expected to be most successful when SB are homogeneous and easily detectable. Secondly, the water predicting performance of the obtained model was assessed when applied on less homogeneous or observable basins (namely heterogeneous excellent and medium SB).

Considering the satisfactory water predicting performances of the model for both excellent and medium SB (Step 2), the model was applied, in a third step, on the whole Sentinel-2 image time series in order to predict the temporal dynamic of water in all the observable SB. This last step ends with a validation of the obtained temporal evolution of water inside the observable SB.

2. SB Dataset Preparation and Characterization of their Detectability

2.1. Delineation of SB

The delineation of SB was completed with the GPS coordinates of the 156 SB recorded in the provincial database and the orthophotoplans using the software *ArcGIS*. The delineation of SB was largely supported by the orthophotoplan from 2018 whereas those from the previous years (1994-2000, 2006-2007, 2009-2010, 2012-2013, 2015, 2016, and 2017) were used to distinguish less easily identifiable SB. Limits of SB were drawn to correspond to the boundaries of the retention infrastructure itself and not the property boundary of the infrastructure site. Assistance of the province has been sought for the rare case where boundaries of the facility were not distinguishable from the property boundaries in order to guarantee an exact delineation of SB. The shapefile resulting from the delineation process comprised 140 polygons and was assumed to be valid for the studied period. Only four SB³ were built between 10 June 2016 and 17 May 2018 (acquisition time of the orthophotoplans). A particular attention was paid to them during the interpretation of the results. The shapefile was re-projected from the projection of orthophotoplans (Belge Lambert 1972) into the coordinates system of Sentinel-2 rasters (UTM 31/ WSG84). Eventually, the information about SB recorded in the provincial platform text file was added to the resulting shapefile.

3. The four SB are named « Buston RER », « BO hirondelles », « BO du merchin » and « Avenue Georgette ».

2.2. Detectability of SB

The delineation of SB highlighted the fact that several SB were hardly observable on orthophotoplans. Therefore, it seemed very likely that Sentinel-2 images would not enable the detection of all the SB of the province. Consequently, a methodology was implemented to qualify, prior to any use of Sentinel-2 images, the expected feasibility to detect each SB based on the adequacy between its characteristics and on the known Sentinel-2 images' limitations (see 2.2.1). Furthermore, a procedure of validation was settled to ensure the accuracy of the attributed qualifications (see 2.2.2).

2.2.1. Assignment of Detectability Level by Photointerpretation

Sentinel-2 surface reflectance images are characterized by the four following major limitations :

- Size of one pixel (100 m²),
- Sensitivity of the signal to obscuring and permanent elements,
- Sensor's point spread function (PSF),
- Residual geometric error.

Taking into account of these four limitations, a methodology was implemented to qualify the ex-ante detectability of SB from Sentinel-2 images. Three levels of detectability (challenging, medium and excellent) were defined according to two variables : the size/shape and the vertical visibility of SB surface.

The variable of size/shape was subdivided into 3 levels according to the number of pixels remaining after the application of a negative buffer of 10 m on each SB polygon (Table 4). The application of a negative buffer on the delineated SB polygon was integrated to the definition of the size/shape variable for two reasons. Firstly, the buffer was settled to limit the impact of the residual geometric error on the detectability of the SB. In fact, the automatic geometric correction of Sentinel-2 images leaves a residual error which causes a slight shift between corrected images. This problematic shift induces lag between the SB depicted on the image and its delimited polygon, leading to a mix of SB and non-SB pixels encompassed within the polygon extent. According to the literature, the current residual geometric error for Sentinel-2 is acknowledged to be at most 0,6 pixel (Pandžic *et al.*, 2016). In our case, the negative buffer of 10 m enabled to offset a maximal geometric error of 1 pixel. With this precaution, an image shift of 1 pixel apart from the center of the polygon would still result in only pixels from the basins intersected by the polygon. Secondly, the buffer was applied to guarantee that the pixel's signals of the buffered SB were not subject to the impact of sensor's point spread function (PSF). PSF of sensor systems show that a significant portion of the recorded signal of each pixel of a satellite image originates from the surroundings of the area represented by that pixel (Huang *et al.*, 2002). When applying a negative buffer of the pixel length, the pixels of the buffered SB should be surrounded by at best one SB pixel. The exact portion is depending on the geometric correction error but in any cases it increases the portion of SB pixels taken into account in the signal. Finally, the size/shape variable was defined to account for the limitation enforced by the size of the Sentinel-2 pixel on the size/shape of identifiable SB. In fact, the variable gives an indication of the size and shape of the basin with the number of pixels remaining after the application of a negative buffer. Intuitively, the number of remaining pixels is expected to be larger for larger basins than for small basins and to be lower for elongated SB in comparison with a SB of same size but more compact shape. This

definition enables the integration of the idea into the variable that bigger and more compact shape SB are allowing a best detectability.

The visibility variable was subdivided into 3 levels according to the percentage of the visible surface (Table 4). Visible surface refers to the area unimpeded by permanent obscuring elements (viaducts, trees, etc.) or their related shadows from azimuth perspective. With this definition, the parameter accounts for the sensitivity of the signal to obscuring elements and reflects the fact that it would be harder to identify SB when their surface can barely be visible from space. Table 4 illustrates the exact definition of these two variables for the three levels of detectability.

TABLE 4 – Definition of the detectability levels (challenging, medium and excellent) for the size/shape variable and the visibility variable. *After the application of a negative buffer of 10m.

	Size/shape	Visibility
Challenging SB	<1 remaining pixel *	<20% surface visible
Medium SB	1<x<9 remaining pixels *	20<x<90% surface visible
Excellent SB	>9 remaining pixels *	>90% surface visible

Each SB was given a level for both variables using a photointerpretation method. The visual interpretation of the size/shape was conducted on a Sentinel-2 image in order to preserve the impact of the Sentinel-2 image's limitations on the detectability of SB. The used image dated from October 10, 2018 and was selected for its cloud free coverage. In contrast, the determination of the visibility has been achieved thanks to orthophotoplans since the presence of obscuring element was not specific to Sentinel-2 images. By offering an increased level of detail in comparison with Sentinel-2 images, orthophotoplans made it possible to better visualize permanent obscuring elements. However, it has to be mentioned that the visibility assessment did not considered the potential gain of SB visibility from the leaf-fall occurring in autumn and winter months, since the acquisition of orthophotoplans dates from spring or summer months. Eventually, the resulting and unique detectability level of each SB was settled to the level of its most constraining variable.

At the end of this stage, SB were qualified by a unique detectability level (challenging, medium and excellent) indicating their ease of identification with Sentinel-2 images.

2.2.2. Validation of the Photointerpretation

Fully aware that subjectivity may potentially have been introduced during the photointerpretation, the validation of the method was performed in order to determine if the assigned levels of detectability were correct.

The validation of the size/shape photointerpretation was assessed by means of a size/shape indicator combined with a shape indicator. Both indicators were calculated on the SB polygons after the application of a negative buffer of 10 m. The shape indicator was defined as the ratio between the area of the polygon delineating the SB and its perimeter. The size/shape indicator was defined by the number of pixels intersecting the polygon, just as the size/shape variable's definition (defined in at the point 3.1). They are distinguished by their calculation method. The size/shape indicator was estimated by the computation of the exact number of pixels intersected by the buffered polygons, whereas the size/shape variable was estimated visually.

Since these indicators were designed to serve the verification of the shape/size level assigned by photointerpretation, they were expected to allow the distinction of three levels related to the different

detectability levels. Challenging SB were expected to be removed after the application of the buffer (number of remaining pixels < 1). Medium SB were expected to count a low number of remaining pixels ($1 < x < 9$) for the size/shape indicator and a low value for the shape indicator. Excellent SB should differ from medium SB with its higher values for the size/shape indicator (number of remaining pixels > 9) and its higher value for the shape indicator.

For each SB, the indicators' values were calculated and confronted to the detectability level of the size/shape assigned by photointerpretation, in order to assess the accuracy of the photointerpretation method.

The validation of the visibility photointerpretation has not been evaluated because the alternatives verification methods (field observation, etc.) were not more objective than the photointerpretation method. Consequently, the ability of the photointerpretation method in assigning correct identifiability levels has only been ensured by the validation of the size/shape photointerpretation. It was assumed that a valid photointerpretation of the size/shape would have resulted in a valid photointerpretation of the visibility, provided that the estimation of the visibility variable is of equal difficulty as the size/shape estimation for a given experimenter.

3. Calibration and Validation of the Model for Water Detection inside SB

The following section describes the steps leading to the calibration⁴ and validation of the classification model for water detection inside SB based on reflectance collected by the three Sentinel-2 calibration/validation images. But before that, the section starts by describing how the reference values used for the model training and validation have been obtained.

3.1. Labelling SB with their corresponding bottom type (Reference values)

Both the training and validation of the model required labelled data since it was chosen to opt for a supervised learning of the classification algorithm.

Since classification model aims to discriminate between water and the other SB bottom types in order to detect water, the labels must account for the bottom types encountered inside SB. The SB delineation with orthophotoplans allowed the identification of four SB bottom types (including water) : vegetation, geomembrane, concrete and water. The latter constituted the four possible labels used to create reference values which represent the state of SB over time.

The labelling of bottom types to the SB was performed by visual interpretation based on Sentinel-2 images. Composite images such as false color, true color and customized composites (Blue, NDWI and AWEIsh) were used to support the distinction among bottom types of SB.

In a first instance, the bottom type corresponding to the time of the image acquisition were assigned to the observable SB (excellent and medium). When there was more than one bottom type, the largest one was assigned to the SB. However, a second methodological choice was to adopt a per-pixel classification approach. Therefore, in a second instance, the unique label of each SB was attributed to the pixels encompassed by their delineated area.

4. The model calibration refer to the determination of the optimal parameters for the training of the model based on trial and error.

3.2. Training and Calibration of the Classification Model on Excellent Basins

The Random Forest (RF) algorithm was chosen to perform the classification of SB bottom types since the literature review revealed its high performance. As mentioned before, a pixel-based classification approach was adopted because the spectral information provided by explanatory variables appeared to be sufficient for the purpose of the study. The pixel-based approach also facilitated the comparison between the images forming the time series. It was also chosen to go for a supervised learning of the algorithm by training it on labelled data. Furthermore, it was initially decided to train the classification on excellent SB as they are the most observable one.

3.2.1. Training

Given these methodological choices, the training approach was to train the classification model on the spectral bands and water indices reflectance belonging to the pixels that encompassed referenced training areas. An extraction per pixel was performed in order to extract these reflectance of the explanatory variables (spectral bands and water indices described at V.2.2.5). The extraction rule states that a pixel was subject to extraction when it intersected the training areas' extent. The dataset resulting from the extraction comprised the pixel reflectance of all the explanatory variables, along with an ID of the training area from which the pixel originated. Eventually, the reference of the training area, assigned to the pixels (3.1.), were added to the previously mentioned dataset. This sequence of operations was repeated for each of the three calibration/validation images.

A random selection of two thirds of the training areas (based on their ID) was applied on the dataset resulting from the extraction process to create the training dataset. Particular attention was paid during the formation of the training dataset, to ensure that the random selection picked SB of various sizes in order to also select two thirds of the pixels. The validation dataset was composed of the remaining third.

The validation of the model consisted of comparing, for each training area, the predicted bottom type with the reference bottom type, both included into the validating dataset. The confusion matrix resulting from the validation process served the evaluation of the classification accuracy with the computation of the overall accuracy and the F1-score metrics (Equations 1 and 4).

3.2.2. Calibration

The parameters of the RF classifier itself were set with optimized values regarding the number of trees and the number of features used at each node. However, the methodology set up to train the classification model (described here before) led to different options regarding the following variables :

- The number and nature of training areas
- The number of calibration/validation images
- The number and the nature of explanatory variables
- The number and the nature of final categories to predict.

The calibration of the model consisted therefore in determining, based on trial and error, which combination of variable's option enabled the training of the most performing model. The following paragraphs described the different possible options for each variable.

The number and nature of training areas are considered among the variables to be calibrated in order

to ensure the representativeness of the training areas. In fact, the way of assigning the pixel's label⁵ implies that the pixels encompassed in heterogeneous SB were attributed the same bottom type irrespective of their differences. Therefore, one can recommend taking into consideration the reduction of the training areas from all excellent SB to homogeneous excellent SB only. Eventually, the addition of additional areas representatives of the bottom types should be considered in order to compensate for the potential reduction of training areas.

The number of calibration/validation images as well as the number and the nature of explanatory variables from which reflectance values are extracted should also be taken into account. Finally, the bottom types categories predicted by the model could be reduced from predicting all bottom types individually to predict only water versus the other.

3.2.3. *Determination of the Most Effective Subset of Features*

Among options to train the model, one relates to the possibility of working with a pertinent subset of explanatory variables. The following methods were tested in order to identify the most discriminating and informative features with the intent to group them in a pertinent subset. All these methods were conducted on a common dataset comprising the pixel reflectance of all the explanatory variables, along with an ID of the training area from which the pixel originated and the reference of each pixel corresponding to the three calibration/validation images.

Graphically

The first method was to graphically assess the discrimination potentials of bottom types for single and pairs of explanatory variables. The discrimination potential of each explanatory variable was depicted by the mean spectral signature of each bottom type (vegetation, geomembrane, concrete and water) and their corresponding probability density of reflectance. The discrimination potential provided by each pair of explanatory variables among the 21 features were represented by scatter plots. The latter represented the reflectance of the first feature of a given pair against the reflectance of the second one. The observations were shaped and colored according to the assigned bottom type of the SB from which they originate at the time of the observation. These graphs were visually analysed in order to identify which single and pairs of explanatory variables were among the most effective at discriminating the bottom types' signature of SB. The most relevant couples as well as single variables were selected as suitable candidates for the subset formation.

Correlation

The second method was to study the correlation among variables in order to prevent the addition into the subset of explanatory variables that are correlated with the ones already forming the subset. The correlation between explanatory variables was quantified in order to identify groups of correlated (Pearson correlation $>0,8$) features. The explanatory variables added to the subset should preferably originate from discrete groups of correlated features.

Anova

The third method was to quantitatively assess the significance of the discriminating power of single feature by means of the statistical test Anova. The discriminating power of a feature was acknowledged to be significant when the corresponding signature of the SB bottom types were significantly

5. The pixel's label were determined by the label assigned to the basin from which they originate (see 3.1.)

different. Only significant feature should be considered to form the subset.

Separability Analysis Using Jeffries-Matusita Distance

The next method sets further the quantification of the discriminating power of the bottom types' signature by considering combination of features (instead of single feature like in Anova). This quantification was performed using the Jeffries-Matusita distance (JM distance) test.

By definition, this test provides a quantitative estimation of the separability between two classes considering observations from a single or multi-variables, based on the probability density function of the two classes. The probability density of classes is assumed to follow Gaussian distribution and are therefore characterized by the following statistical parameters : the mean vector and the variance-covariance matrix. These parameters serve the quantification of the potential overlapping area in relation with the classes separability. It is supposed that the smaller (bigger) the overlapping area, the best (worst) the separability (Thomas *et al.* 1987). The calculated JM distance is asymptotic to 2 when signatures are completely different, and tends to 0 when signatures are identical. Its mathematical definition is presented here after.

$$JM = 2 \times (1 - e^{-B}) \quad (7)$$

where

$$B = \frac{1}{8}D^2 + \frac{1}{2} \ln \left(\left| \frac{\sum_j + \sum_k}{2} \right| / \sqrt{\left| \sum_j \right| \left| \sum_k \right|} \right) \quad (8)$$

and

$$D^2 = (\mu_j - \mu_k)^T \left(\frac{\sum_j + \sum_k}{2} \right)^{-1} (\mu_j - \mu_k) \quad (9)$$

where μ_j and μ_k correspond to the expected reflectance of each class respectively, and $\sum_j$ and $\sum_k$ are unbiased estimates of covariances (Wardlow *et al.*, 2007)

In essence, this test only provides estimation of the separability between pairs of classes for a given number of features tested (Thomas *et al.*, 1987). A choice was made to quantify the separability of all classes with the maximal JM distance separating the pair of classes.

It is noteworthy to mention that the evaluation using the JM distance requires the singularity of the matrix, since its definition comprised a matrix inversion. Consequently, the combinations to be tested were limited to the combinations of features constituting a singular matrix. Nevertheless, the number of combinations to be tested remained very large. For those two reasons, the Sequential Forward Floating Selection (SFFS) algorithm was used to automatically generate subsets of features from the original feature space and compare these feature subsets in terms of discrimination using the Jeffries-Matusita distance test as an evaluation function. The search algorithm and the evaluation function were already implemented in the "VarSel" package provided by R under the names 'VarSelSFFS' and 'JMdist' respectively.

For each date, the SFFS algorithm was used to determine the selection of features and the resulting maximal JM distance of the most discriminating subset predicted when the complete feature dataset was given as an input to the algorithm. The resulting maximal JM distance was then compared to the one obtained when the input dataset was reduced from the full set of features to the following subset

(b2, b4, b8A, b11, NDWI, AWEInsh, AWEIsh, LWSI and s). The features forming this subset were selected according to the outcomes of precedent methods. It is foreseen that this subset should allow a reduction of the number of features while maintaining sufficient JM distance. Complementary features (b8, b7 and mNDWI) were successively added to the subset and fed as an input to the algorithm in order to assess their impact on the resulting maximal JM distance.

Gini Index of Random Forests Classification

A last option for evaluating the discriminating power of features combination was to perform the classification and subsequently analyses the importance of the feature used. The Gini Index was used as an indicator of the importance of the feature for the whole classification. Since each evaluation of the feature importance requires a classification, this method is more limited regarding the number of features combinations tested than the SFFS-JM distance algorithm. The Gini Index was calculated for both the complete and subset input dataset (retained from the SFFS-JM distance algorithm) considering the three calibration/validation images, both individually and as a group.

3.3. Validation of the Classification Model

The last operation of the model development (Step 2) was to assess its predictive capability when it is applied on basins type or images that did not serve the training. Since the calibrated classification model (Chapter VII.2.2.2) was trained on homogeneous excellent with the September image only, this section has assessed the performance of the model when applied on the homogeneous excellent, on the heterogeneous excellent and on the medium SB considering the three calibration/validation images.

For all tested types of SB, the model performance was assessed by means of a traditional comparison between the predicted bottom type with the corresponding reference for each pixel. A comparison of the most frequently predicted bottom type and the corresponding reference at the level of SB was also performed. For both approaches, predicted bottom types were obtained by applying the retained model on the reflectance extracted from the SB's pixel for the three calibration/validation images. In addition to these quantitative approaches, a visual comparison of the initial Sentinel-2 image with its classification was conducted in order to spatially assess the classification validity.

4. Water Monitoring inside SB

This section is dedicated to the last step of the methodology implemented to achieve the monitoring of the water inside SB. More specifically, the section starts by describing how the model is applied on the whole Sentinel-2 time series in order to predict the temporal dynamic of water in all the observable SB. This is followed by a presentation of the validation of the obtained temporal evolution of water inside considered SB.

4.1. Application of the Classification Model on TS for both medium and Excellent SB

Considering the satisfactory predicting performance of the model for both excellent and medium SB, the calibrated classification algorithm was applied on the whole Sentinel-2 image time series in order to predict the temporal dynamic of water for each of those basins' types.

In order to do this, the pixels reflectance values were extracted from each date of the time series when they encompassed the polygons corresponding to the type of SB considered. The extracted pixel

reflectance were then gathered in a dataset per date of the time series. Each dataset was composed of samples characterized by the extracted reflectance values for all features and the ID of the SB from which pixels originate, but they had no value regarding their attributed bottom type. The number of samples was varying with the tile and the type of SB considered (Table 5).

TABLE 5 – Number of samples used for classification on medium and excellent SB types in each tile UES or UFS

	Tile UES	Tile UFS
Level Medium	166	656
Level Excellent	355	1216

The calibrated classification model was then applied successively to each dataset containing the pixel reflectance values and SB ID, in order to complete the dataset with the bottom type prediction for each pixel. Further on, these datasets were rearranged in order to group the reflectance values and their corresponding bottom type prediction per basin for all the dates of the time series, and no longer per date for all the basins. Finally, the proportion of water classified pixels was computed per basin for each date of the time series. This resulted in the desired temporal evolution of the proportion of water in the basins. In complement, the computation of the water surface (*fwater*) was calculated in order to provide information about the surface encompassed in that proportion measurement.

The sequence detailed here before was applied in an identical way for excellent and medium basins and was repeated twice for each SB type : one time for UES tile and another time for the UFS tile. Only the shapefile used for the extraction process was adapted to the specific situation.

4.2. Validation of the Temporal Evolutions

The validation of the temporal evolutions resulting from the third step was performed according to two methods.

The first one was based on a visual interpretation of the adequacy between the initial Sentinel-2 images and its classification. This method was divided into two approaches in order to assess simultaneously the temporal and spatial dimension. On the one side, the “temporal” approach consisted in visually assessing the accuracy of the classification for a sample of 10 SB of each type (medium and excellent) randomly selected on all the classified Sentinel-2 images from the time series (around 50 images per SB). The visual assessment was performed through a comparison of the initial Sentinel-2 image, the classified one and the attributed water proportion for each date of the time series. To save time and effort, this has been done with automatically generated Graphics Interchange Format (GIF). On the other side, the “spatial” approach was to visually assess the accuracy of the classification for a sample of 10 Sentinel-2 images on all the medium and excellent SB. The 10 images were randomly selected among a restricted group of images that fulfilled the three following conditions : 1) covering both tiles at the same time, 2) having small cloud coverage and 3) being evenly distributed across the temporal coverage (end of 2016, 2017 and 2018). The first two conditions were settled to assess the maximum of SB at a given date and the last one ensured the diversity of encountered image.

The second method consisted in comparing the evolution of water proportion obtained by classification with the daily water-levels recorded by lysimeters from the Aqualim network.

Chapter VII. Results

This chapter is dedicated to the description and the display of the different experiment's results. This chapter follows a structure organized in three sections (Sections 1, 2 and 3) according to the three pursued sub-objectives.

1. SB Dataset Characterization

1.1. Delineation of SB

The boundary delineation was successful for 140 SB out of the 156 artificial open basins originally recorded in the provincial dataset. The 16 other SB recorded in the database was not be found for the following reasons :

- Three SB had already been recorded in the database while they were still at a project stage or under construction (as correctly mentioned in the database field called « operating status »).
- Two SB were still acknowledged to be at the project stage while they had been built and were deconstructed ever since.
- Three SB were impossible to locate even though they were recorded in the platform. For them, the provincial administration in charge of the platform update sought for clarifications from the administrations who reported those basins in order to verify their existence. At the time of the delineation, the updated information about the location of the three basins was not yet reported by the concerned authorities.
- Four SB were recorded in the platform and the province confirmed their existence but could not be identified from orthophotoplans.
- Four other SB were impossible to delineate owing to the presence of trees.

This step also helped to identify and revise some outdated information of the provincial platform database. In fact, the inspection of orthophotoplans revealed that two SB were still characterized as an ongoing project even though they were constructed. However, the real situation can not always be depicted so easily by the orthophotoplans. As an illustration, two existing water bodies were characterized as an ongoing project. This information seemed to be outdated since the water bodies were visible, but it was not because the water bodies were in the process to be transformed in man-made SB. Finally, a new basin has been located and reported at the province administration for recording. The annexe B lists the exact name of each SB referred to in this section.

1.2. Detectability of SB

The 140 delineated basins were given a detectability level (excellent, medium or challenging) to qualify ex-ante the feasibility of their identification with Sentinel-2 images. The assigned level has been determined by the photointerpretation of two variables, accounting for the size/shape of the basin and for the visibility of its surface respectively. The results obtained by photointerpretation for the parameter size/shape were further confronted to two quantitative indicators in order to assess the validity of the photointerpretation method. The visibility parameter was not considered in this assessment because of the lack of quantitative index.

This sub-section starts with the outcomes of the assessment process and pursues on the levels of detectability resulting from the validated photointerpretation method.

1.2.1. Validity Assessment Outcomes

Primary Results

The visual photointerpretation of the size/shape variable resulted in the following distribution of detectability levels : 28 SB identified as challenging, 47 SB identified as medium and 65 SB identified as excellent. The following paragraphs present the results obtained from the validation by objective indicators.

The buffer was expected to discriminate challenging SB by eliminating them as they theoretically are characterized by less than 1 pixel after the buffer. In total, 35 basins were removed following the application of a negative buffer of 10 m. As predicted, a majority of the removed SB were part of the challenging category (24 out of 35). The eleven other SB were also deleted even though they were not belonging to this category. Ten SB were part of the medium category and one basin from the excellent category. Consequently, the application of the buffer highlighted 24 cases of good prediction and 11 cases of overestimation resulting from the photointerpretation. This statement was compatible for both indicators since they both rely on the buffer.

The shape indicator was expected to discriminate between medium and excellent SB with an increased value for increasing size/shape levels (from challenging to excellent). The expected trend of indicator values was confirmed but the ranges corresponding to the identifiability levels overlapped. This problematic situation prevented the discrimination between the detectability levels by the size/shape indicator. At best, the indicator allowed a partial differentiation between challenging-medium and medium-excellent level.

The size/shape indicator enabled a precise discrimination between medium and excellent SB according to the number of pixels intersecting the buffered polygon. This indicator revealed that 35 out of 47 medium SB and 39 out of 69 excellent SB were correctly classified by the photointerpretation. However, this indicator also highlighted incorrect assignments by photointerpretation. Four challenging SB counted more than one remaining pixels after the buffer and therefore should have been classified in the medium category according to the size/shape indicator. Similarly, two medium SB counted more remaining pixels after the buffer than its definition warrants and thus should be classified as excellent according to the size/shape indicator. Finally, 25 other excellent basins were considered as misclassified by the indicator because of their lower number of remaining pixels after the buffer in comparison with their definition.

To sum up, 70 percent of the size/shape levels assigned by the photointerpretation method were confirmed to be correct by the size/shape indicator and the buffer. Table 6 summarizes the number of well and misclassified basins according to the size/shape indicator and the buffer.

TABLE 6 – Contingence matrix for the comparison between the levels assigned by photointerpretation with the one resulting from the quantitative size/shape indicator

		Definition of the size/shape indicator		
		Level Challenging	Level Medium	Level Excellent
Photointerpretation	Level Challenging	24	4	0
	Level Medium	10	35	2
	Level Excellent	1	25	39

Revaluation

The photointerpretation and the quantitative approach converged towards the same size/shape identifiability levels in 70 percent of cases. The other 30% controversial basins were reassessed by photointerpretation in order to decide which of the two approaches suits best to the reality. The revaluation outlined that :

- The SB removed after the application of the buffer should be classified in the challenging category irrespectively of their initial class.
- The four controversial challenging SB should not be upgraded to the medium category as suggested by the quantitative approach because the threshold of 1 remaining pixel was reached while counting pixels that originated from a minor rest of the buffered polygons (Figure 9.a & b).
- The two controversial medium SB should also not be upgraded to the excellent category as suggested by the quantitative approach because the threshold of 9 remaining pixels was reached while counting pixels that originated from a minor rest of the buffered polygons (Figure 9.c).
- Among the 25 excellent SB claimed to be overestimated, the revaluation confirmed the overestimation by photointerpretation for 18 of them and were downgraded to the medium level. The 7 left were sufficient to be maintained to the excellent level.

Table 7 summarizes the number of well and miss-classified basins according to the size/shape indicator and the buffer after revaluation of the controversial SB.

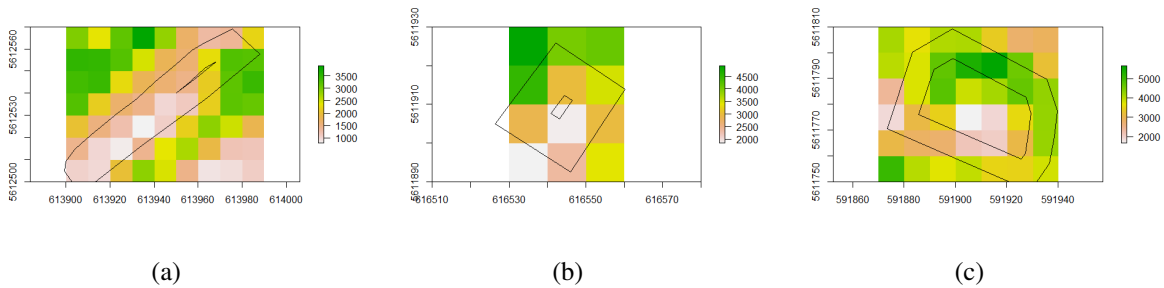


FIGURE 9 – Controversial challenging (a,b) and medium (c) SB that were not upgraded to the higher level as suggested by the quantitative approach because the the threshold corresponding to the higher level was reached thanks to remaining pixel originating from a minor rest of the buffered polygons.

TABLE 7 – Accuracy of the photointerpretation by the comparison between the levels assigned by photointerpretation with the one resulting from the quantitative size/shape indicator, after revaluation.

		Definition of the size/shape indicator		
		Level Challenging	Level Medium	Level Excellent
Photointerpretation	Level Challenging	35	4	0
	Level Medium	0	55	2
	Level Excellent	0	5	39

After revaluation, the photointerpretation and the measurement approach relying on quantitative indicators led to the same size/shape identifiability levels in 90 percent of cases.

An agreement rate of 100 percent between the photointerpretation and the quantitative approach related to the size/shape variable could have been achieved with a strict compliance of the quantitative indicators. This would further require putting the observation of the photointerpretation aside. Since

the revaluation outlined that the reality cannot be fully depicted by the quantitative approach, this strict compliance with the indicators is not recommended.

Therefore, the reevaluated identifiability levels assigned to the size/shape parameter were eventually acknowledged to provide the best results. A revaluation was also performed for the identifiability levels assigned for the visibility parameter but no modification was required. The resulting identifiability levels will serve the rest of this thesis and are presented in the following subsection.

1.2.2. SB Detectability Levels

This section presents the detectability levels of SB resulting from the validated photointerpretation method. Among the 140 considered SB, 31 SB were assigned an excellent detectability level meaning that both their size/shape and the visibility of their surface were very convenient for their identification with Sentinel-2 images. A medium detectability level was allocated to 54 SB. These medium basins were at least characterized by a variable of size/shape or of visibility of the surface that restrains their detectability with Sentinel-2 images. Finally, 55 SB were given a challenging detectability level when their detectability with Sentinel-2 images was highly compromised by at least one of its variable of size/shape or visibility. As a result, the proportion of SB classified with an excellent, medium or challenging detectability was respectively 22%, 38% and 40%. This indicates that a majority of SB is potentially observable using Sentinel-2 images but the proportion of challenging basins is not negligible. Figure 10 represents the homogeneous distribution of detectability levels across the study area and can also be seen as a map of potentially monitorable SB (medium and excellent) versus potentially non-monitorable SB (challenging). From this map, one can also deduce that the area potentially monitorable SB accounts for 81,9% (370 220 m²) of the total area covered by SB (451 567 m²).

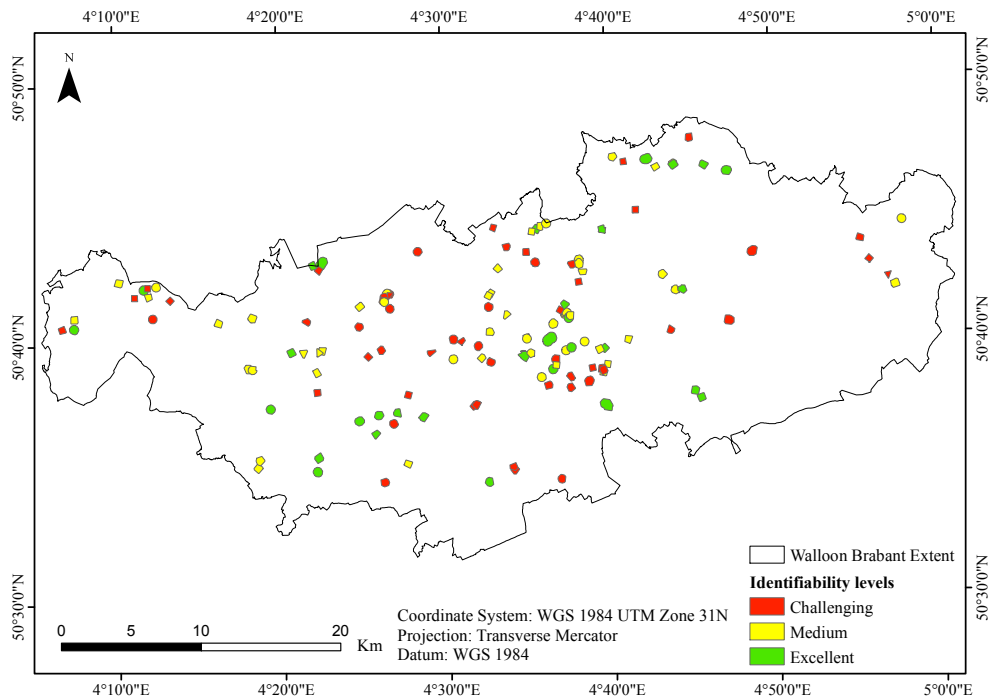


FIGURE 10 – Spatial distribution of detectability levels assigned to SB

When considering the levels of each variable, SB can be categorized into seven groups :

1. Challenging visibility and excellent size/shape (1-3) [6/140]
2. Medium visibility and challenging size/shape (2-1) [12/140]

3. Medium visibility and medium size/shape (2-2) [22/140]
4. Medium visibility and excellent size/shape (2-3) [9/140]
5. Excellent visibility and challenging size/shape (3-1) [15/140]
6. Excellent visibility and medium size/shape (3-2) [23/140]
7. Excellent visibility and excellent size/shape (3-3) [31/140]

Their respective proportion (specified in brackets) indicate a balanced distribution of the SB in the groups. Examples of cases 7, 6 and 9 are illustrated in Figure 11.

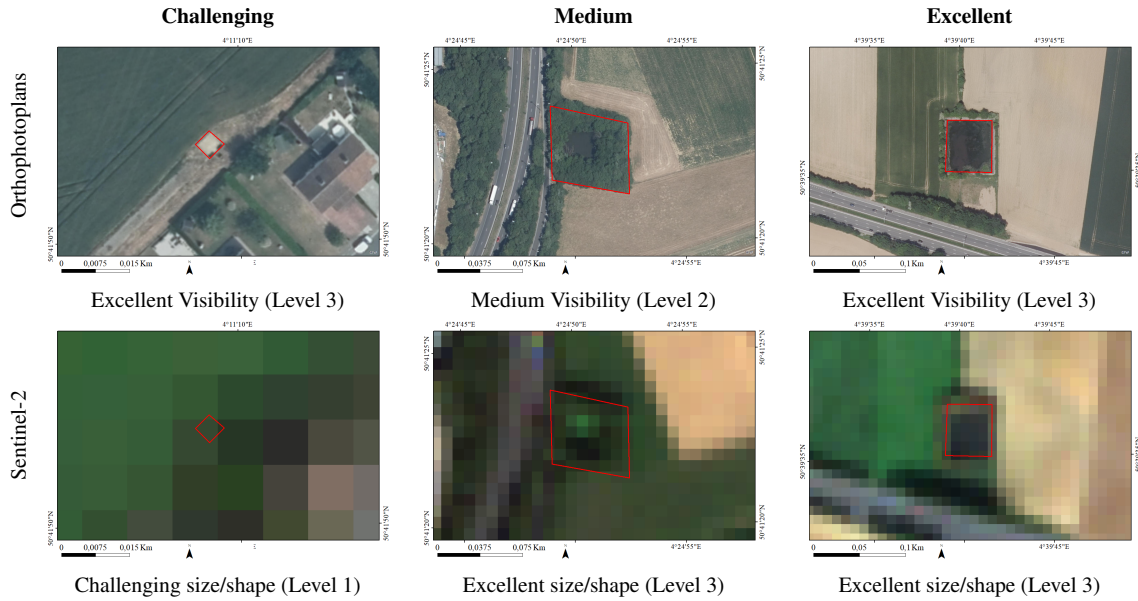


FIGURE 11 – Three detectability levels of SB by rows : 1) challenging, 2) medium and 3) excellent. Two different data are shown by lines : 1) orthophotoplans and 2) Sentinel-2 to represent the detectability sub-levels relative to the visibility and the size/shape of the SB respectively

Regarding the most constraining variable, it was found that :

- 50 SB were characterized by lower values for their size/shape variable than for their Visibility variable.
- 28 SB were characterized by higher values for their size/shape variable than for their Visibility variable.
- 62 SB were characterized by equal values for their size/shape variable and for their Visibility variable.

This points out that the two variables are evenly constraining for a majority of SB. When there is a difference, the size/shape variable tends to be the most constraining one.

1.3. Conclusion

From the SB dataset characterization, one can conclude that a majority (140/156) of SB, initially recorded in the provincial platform of Walloon Brabant, were observable and delineated by means of orthophotoplans. Furthermore, the ex-ante characterization of the SB detectability suggested that a majority (85/140) of SB was potentially observable from Sentinel-2 images, even though the proportion of basins with a challenging detectability is not negligible. The attempt of monitoring will therefore be conducted on the 85 observable (medium and excellent) SB.

2. Calibration and Validation of the Classification Model on Excellent SB

2.1. Bottom Types Labels

To create the labelled reference values, a visual interpretation of the bottom types encountered in the SB was performed based on different composites Sentinel-2 images at the three dates.

As a result, false color composites enabled the distinction between 3 types of bottom (water, vegetation and concrete) out of the desired four (water, geomembrane, vegetation and concrete). In fact, a distinction was drawn between water, vegetation and concrete bottom types through the proposed method thanks to their distinct reflectance values in the canals Green, Red and NIR (Figure 12). Water appeared in black (low reflectance in the three bands), concrete and urban areas were characterized by a turquoise colour (high reflectance in the Green bands and low to medium in the Red bands) and vegetation was revealed by a pronounced red colour (high reflectance in the NIR bands and low to medium in the Red and Green bands).

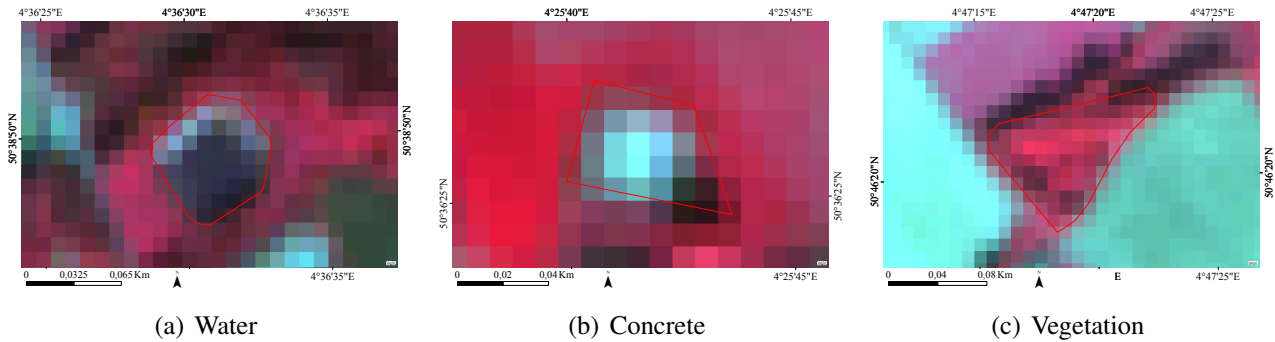


FIGURE 12 – Bottom type discrimination using the Sentinel-2 false-color composite acquired on October 10, 2018

On the contrary, the geomembrane bottom type, distinct from water with orthophotoplans thanks to its brightness (Figure 13), could not be differentiated from water based on Sentinel-2 false-color composites, as they both showed up in black. Additional color composites, such as true-color composites and a Band 2 (Blue), NDWI and AWEIsh composites, were tested in order to address the challenge of distinguishing between geomembrane and water bottom types. In addition, a graphical analysis of the bottom types' reflectance was performed (as described at 2.2.1.) to confirm or refute the conclusion resulting from the visual interpretation of composites. These investigations were performed in both a geomembrane at the bottom of a SB and a non-SB related geomembrane. The latter was found to be exposed during the redesign work of the engineered landfill located at the Cour-au-Bois site. Its size and visibility suited the assessment of its discriminating capacity between the geomembrane and water bottom types. Performing this assessment with a non-related geomembrane ensured that the difficulty of distinction between those types was not due to the presence of water above the geomembrane. In fact, contrary to the SB geomembrane, it is very unlikely that the non-SB related geomembrane was covered by water by the time of the acquisition date (10/10/18).

Unfortunately, none of the tested composites enabled the discrimination between geomembrane and water, as shown in Figure 13. This conclusion was confirmed by the graphical analysis of the spectral signature of bottom types (done at section 2.2.1.).

As the distinction of the geomembrane from the water bottom type has not been possible, the geomembrane bottom type was merged with the water bottom type and resulted in the category «water-

geomembrane». Concrete and vegetation bottom types were maintained in distinct categories. Consequently, the basins were attributed three bottom types : water-geomembrane, vegetation and concrete over the three dates.

The consequences of this merger are expected to be minor as it only compromised the monitoring of the two observable SB (out of 140) which display a geomembrane bottom type. The monitoring of the other is not impacted by this fusion since water is the sole bottom type that is susceptible to be added across time. Consequently, for the sake of simplification, the term 'monitoring of water' will be used to refer to the monitoring of geomembrane-water.

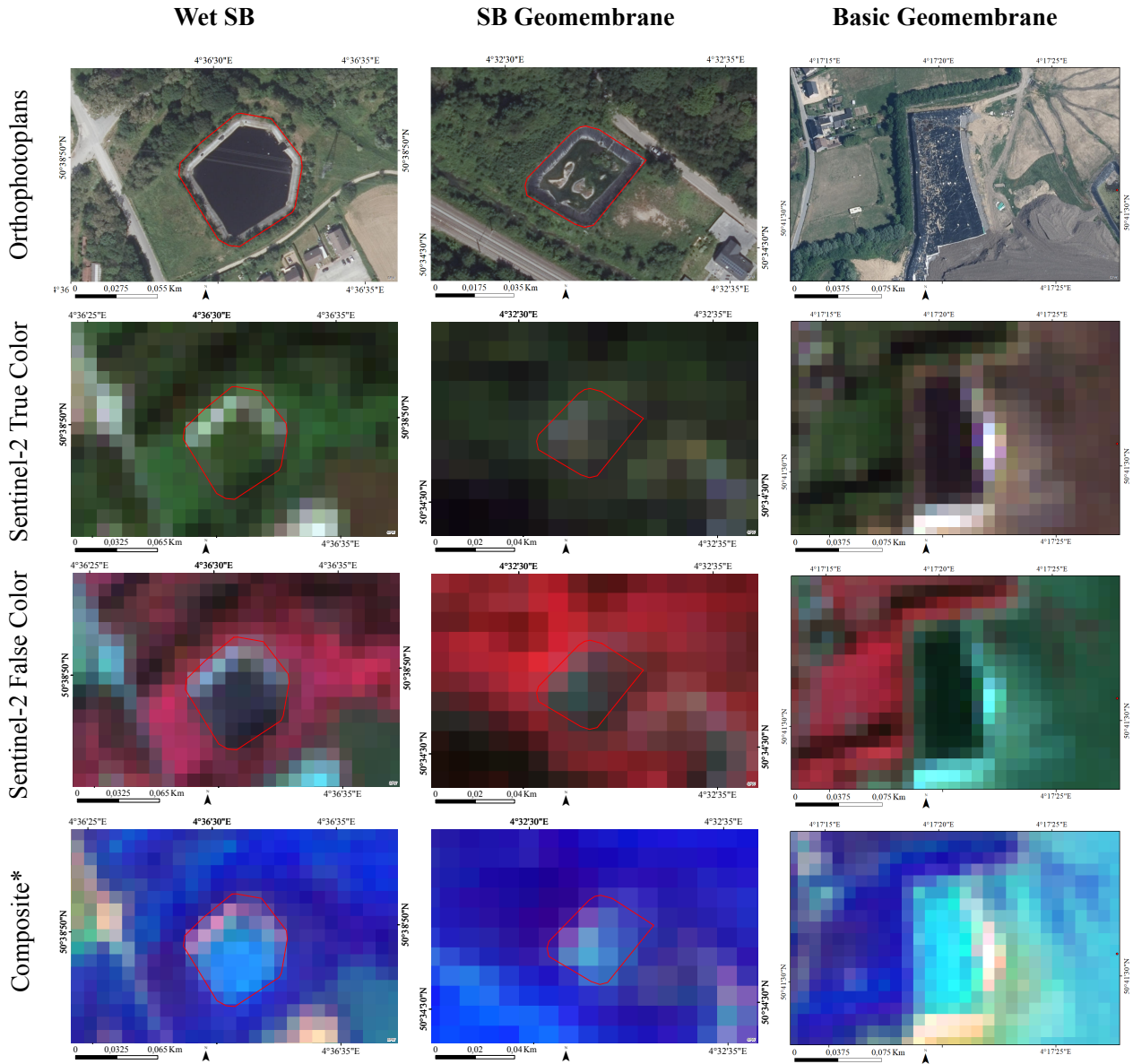


FIGURE 13 – Discrimination between water and geomembrane bottom types, presented by rows : 1) water inside SB, 2) geomembrane inside SB and 3) geomembrane outside SB. Four different Sentinel-2 composites are shown by lines : 1) false-color, 2) true-color and 3) Band 2 (Blue), NDWI and AWEIsh composite acquired on October 10, 2018.

2.2. Calibration of the Classification Model

2.2.1. Determination of the Subset of Features

Among options to train the classification model, one relates to the possibility of working with a pertinent subset of features, in order to reduce the number of features to be calculated and thus the computation effort requirements, while achieving of a comparable or even higher accuracy. Different methods were tested in order to bring to light features that were among the most discriminating ones in the three calibration/validation images.

The following section reports the results of each method while considering their temporal variability. A comparison of the different methods' outcomes is then performed in order to highlight an effective and time-consistent subset of features.

Graphically

The first method was to assess visually the discrimination potential of bottom types for single and pairs of feature. The graphical analysis of the single feature's potential was supported by two graphs : the mean spectral signature of the excellent SB bottom types (Figure 14) and the probability density of their reflectance (Figure 15). These graphs enable the categorization of features into three categories depending on their ability to separate bottom types :

- Poor : mNDWI, mLSWI, MSI, SWM, h, LSWI*, NDWI*
- Medium : Band 2*, Band 3, Band 4, Band 5, s*
- Excellent : AWEIsh, AWEInsh, Band 6, Band 7, Band 8A, Band 8-v, Band 11, Band 12

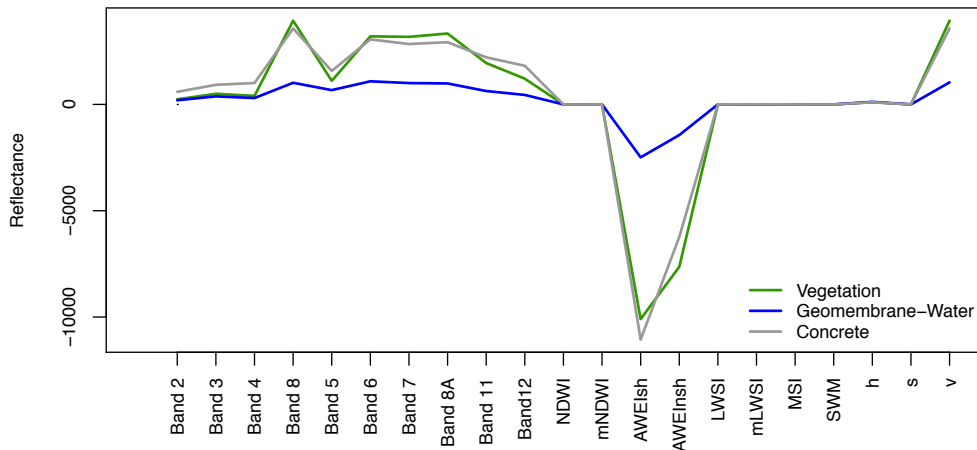


FIGURE 14 – Mean spectral signature of the excellent SBs' bottom types (vegetation, water-geomembrane and concrete) for each explanatory variable (bands and indices), coming from the Sentinel-2 image acquired on September 18, 2018.

Through this analysis, it was noticed firstly that mLSWI expressed an aberrant distribution. Secondly, the index v appeared to follow exactly the same trend as the Band 8. This is not consistent with its definition (Table 3) but it indicates that v does not provide additional information on the Band 8.

A temporal variation was observed on the distributions of reflectance probability but not in the mean spectral signature graphs. The features marked by an asterisk were expressing significantly different probability density distribution in time and thus were categorized in the medium or in the poor groups depending on the considered image.

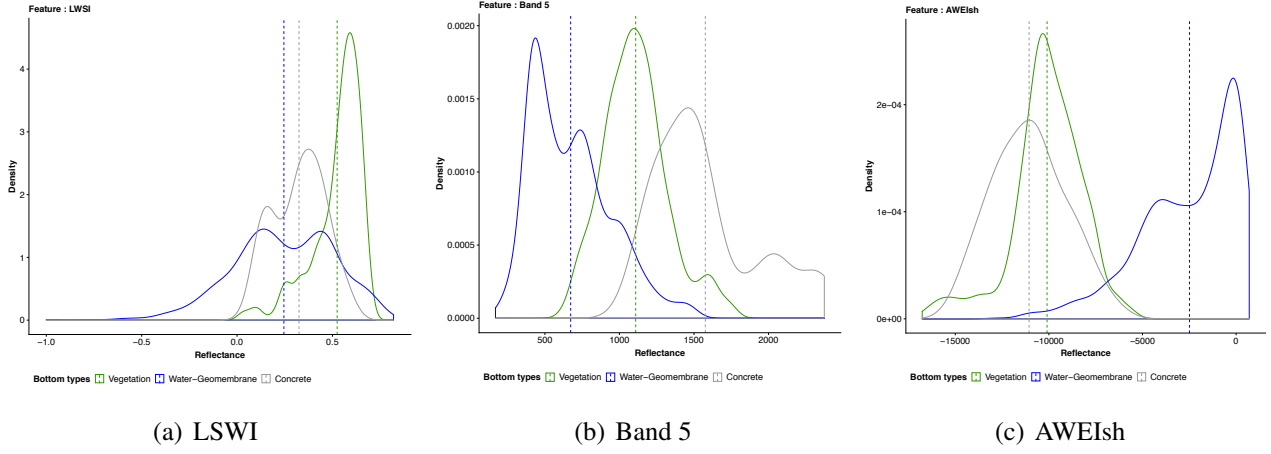


FIGURE 15 – Illustration of the probability density of bottom types' reflectance for poor, medium and excellent discriminating features, more precisely with LSWI, Band 5 and AWEIsh.

The assessment of bottom types' discrimination potential resulting from pairs of features was performed by means of scatter plots. The latter illustrate the reflectance of all pixels belonging to excellent SB, observed in features of each pair (Figure 16).

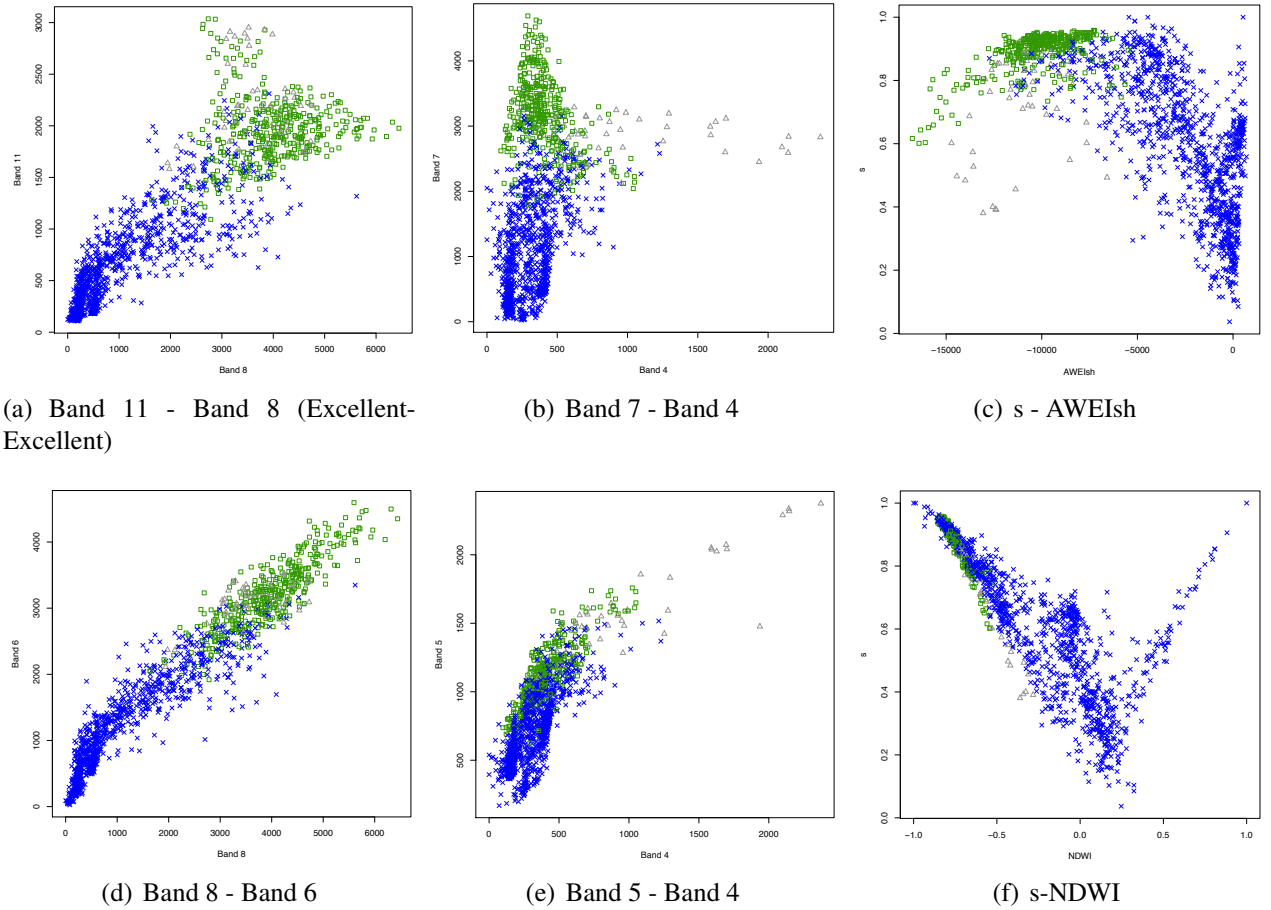


FIGURE 16 – Scatter plot of pixels' reflectance for pairs of excellent features (correlated (d) or not (a)), for pairs of medium and excellent features (b), for pairs of medium features (e), for pairs of excellent and exceptional mediocre features (c) as well as for excellent and basic mediocre features (f). The bottom types are distinguished according to the following colours : "Water-Geomembrane" in blue, "Vegetation" in green and "Concrete" in Gray. Observations from the September image.

These scatter plots demonstrated firstly that pairs of excellent features were providing a high resulting separability between bottom types, as long as the features constituting the pair were not correlated (Figure 16.a). In fact, pairs of correlated features expressed very limited potential of discrimination (Figure 16.d). Secondly, features with excellent discriminating power have proved to enhance the bottom type separability of the pair when they are combined with less discriminating features (medium or mediocre). An example of such improvement in a medium-excellent pair is illustrated in Figure 15.b. However, in some cases, the improvement provided by the presence of an excellent feature in the pair does not suffice to make the combination mediocre-excellent usable (Figure 16.f). Only exceptional mediocre features (Band 2, NDWI, LSWI and s indices) gave surprisingly excellent discriminating power when they are combined with features recognized excellent (Figure 16.c). The impact of so-called exceptional mediocre features can be explained by the different order of the bottom types between their distributions in comparison with the ones of excellent features. As an illustration, the distribution of concrete for the LSWI feature (Figure 15.a) is centered on the opposite side to the vegetation with the water-geomembrane bottom types in the middle, whereas the concrete was comprised between the vegetation and water-geomembrane bottom types. Therefore, the combination LSWI-AWEIsh enables the differentiation of concrete from the other bottom types (Figure 16.c). Furthermore, the pairs of medium features appeared not to provide a sufficient discrimination (Figure 15.e). The same was observed for pairs of mediocre features. This can be caused either by their limited discriminating potential or by their correlation.

Regarding the temporal variation of the discriminating power, it has been noticed that the observations vary in time but that do not modify the aforementioned elements. This only renders certain combination more pertinent at certain dates.

Consequently, the outcomes from the graphical analysis of pairs of features indicate that combinations of excellent-excellent, excellent-medium and excellent-exceptional mediocre features provide the best discriminating performance and that the correlation should be avoided when possible.

Anova

The second method was to evaluate quantitatively the individual significance of the features in describing the bottom types by means of the statistical test ANOVA. Apart from the mLSWI index ($0.9996 > 0.05$), all the features were proved to be very significant for a confidence level of 95% ($2.2e-16 < 0.05$).

Correlation

The third method was to highlight correlated bands among the full feature dataset. As expected, the correlation matrix (Figure 17.a) revealed correlated features among a spectrum of features correlation. Pairs of features were acknowledged to be correlated when their Pearson coefficient exceeded 85%. Correlated features were highlighted in Figure 17.b. This figure enabled the distinction of the following groups of correlated features :

- Band 3, Band 4
- Band 6 , Band 7, Band 8, Band 8A, NDWI, AWEInsh, AWEIsh, v
- NDWI, mNDWI, SWM*

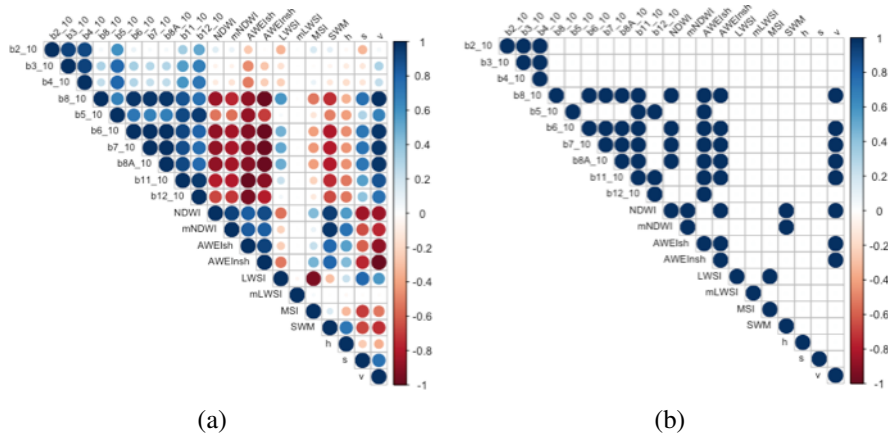


FIGURE 17 – Correlation matrix of the explanatory variables (Sentinel-2 bands and retained water indices), based on reflectance measurements extracted from the 18/09/2018 Sentinel-2 image.

These groups were constant over time, with the exception of the SWM index (marked with an asterisk) which was not always exceeding the threshold of 85%.

Sequential Forward Floating Selection Using Jeffries-Matusita Distance

The fourth method was supported by the Sequential Forward Floating Selection (SFFS) algorithm which enabled the selection subset of features with the best resultant discriminating power according to the evaluation function of Jeffries-Matusita distance.

This iterative algorithm selects an increasing number of features from the input dataset and computes the JM distance at each step. The features to be primarily selected are characterized by an individual high class's separability and thus, high JM distance. The final subset is obtained when the addition of features does not provide a significant increase of the resulting distance (has reached the maximal value) or impede the singularity of the matrix.

Regarding the methodology applied, a selection of features and the resulting discriminating power were evaluated five times : first with the full set of features as an input, second with an input dataset voluntarily reduced to nine features and the last times by successively adding the following features : Band 8, Band 7, mNWSI, to the reduced input dataset. The reduced input dataset was formed according to the results obtained from the previous methods and comprised the following features : Band 2, Band 4, Band 8A, Band 11, NDWI, AWEInsh, AWEIsh, LWSI and s. Theses five assessments were conducted on the data extracted from the calibration/validation images respectively, as well as considering the three images at the same time.

By letting the algorithm freely determine the best subset of feature among the full set fed as an input, it was observed that the individual discriminating power of the features (and their related JM distance) varied over time, thereby changed the order of importance of features among the considered dates. Since the algorithm performs its selection of features based on this order, the resulting subsets were composed of features differing considerably with regard to their nature and number, depending on the date considered (Table C. 1, Figure 18). Even though they differ in terms of the nature and number of selected features, retained subsets tended towards a similar discriminating potential ($\approx 1,4$) for all the investigated dates (Figure). This suggested that 1,4 was the maximal JM distance achievable and that there were many combinations of features to achieve this goal accuracy.

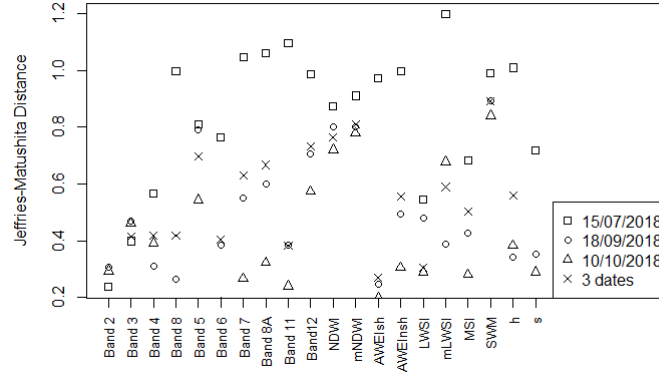


FIGURE 18 – Variability of the Jeffries-Matushita distance for each feature for the considered calibration/validation images as well as considering simultaneously the three dates

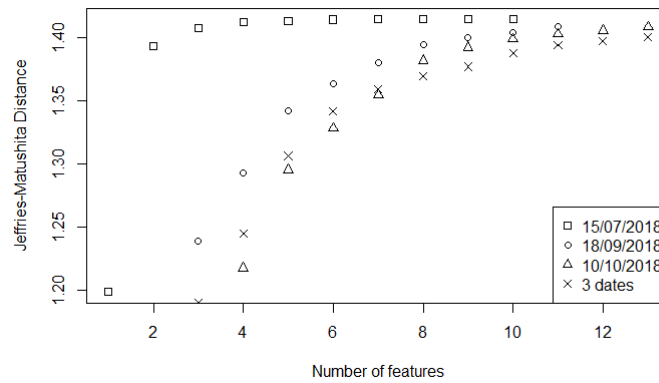


FIGURE 19 – Jeffries-Matushita distance achieved with increasing number of features selected by the SFFS algorithm

By forcing the algorithm to select features among the reduced input dataset, it was observed that the reduction of the features in the input dataset resulted in a diminution of the JM distance, ranging from 2,83 to 5,12% (Table C. 2). The majority of selected features were common to the different dates. The difference between the final subsets lied in which of the two features Band 2 or s were selected. This indicated that they were the two less discriminating bands of the nine features forming the reduced dataset and that the worst of the two bands was not consistent in time.

The addition of Band 8 has revealed a preferential selection of the Band 8 over one of the Band 2, Band 4 and s depending on the date. This has resulted in both a rearrangement of the features and an increased of the JM distance of 1,2% in comparison with the previous case (Table C. 3). After the addition of Band 7 in the reduced input dataset, it was observed that the band 7 has been selected in a unique case out of 4. This has resulted in a decreased JM distance (-3,1%) in the unique case where Band 7 has been selected; in the other cases, the JM distance increased (0,6%) in comparison with the previous cases because the non-selection of band 7 has enabled the selection of a more discriminating bands (B2, B4 or s depending on the dates) (Table C. 4). Eventually, the addition of mNDWI in the reduced input dataset resulted in a selection of NDWI in all cases. An average increased JM distance (1,397) to a level almost equivalent to the maximal JM distance (1,407) was observed following its addition to the input reduced dataset (Table C. 5).

To sum up, the results demonstrated that the algorithm fails to find an optimal subset which is common to all the considered dates with the full input dataset, but the obtained best subsets all tended towards a similar maximal separability (JM distance $\approx 1,4$). Consequently, the most pertinent subset needs to

be stable in time and tends to a JM distance of 1,4 (with a limited number of features). The results obtained while limiting the input dataset to nine features showed that reducing the number of input variables led to a decreased JM distance of the optimal subset in comparison with the one obtained from the full set of features. The addition of features that are more informative bands than the ones in the reduced input dataset (such as Band 8, mNWDI) allows to offset this decrease while achieving a JM distance equivalent to the one obtained with the full dataset. The addition of less informative band than the one initially in the reduced input dataset (Band 7) results in a decreased JM distance when selected. If the band is not selected, one other band could be chosen in its place and its impact would be dictated by its discriminating power relatively to the one already in the reduced input dataset.

Finally, a choice was made to opt for the reduced subset to which Band 8 was added because this set is stable over time. This is of great importance when dealing with time series. Furthermore, this set enables the reduction of input features from 21 to 9 while achieving separability comparable to the ones obtained with the full input dataset (1,394 vs 1,407).

Random Forest Gini Index

The last method was to assess the discriminating power of features used by the RF algorithm during a classification. The results (illustrated in C. 1) confirmed that the features importance rankings resulting from the application of the RF algorithm on both the full and reduced set of features were variable among calibration/validation image. Furthermore, with the full set of features, a plateau in the decrease of gini index is reached before having utilized all the features. With the reduced set of features, a comparable level of mean Gini decrease is obtained without stability zone which suggests that the subset contains only essential features.

Conclusion

In conclusion, the visual analysis of the bottom type discrimination for single and pairs of features highlighted interesting candidates for the subset. This analysis also demonstrated the undesirable presence of correlation among bands and the absurd behaviour of mLSWI. In fact, mLSWI was the only feature to be non-significant according to the statistical test ANOVA. The correlation analysis further defined three groups of correlated bands. Finally, a subset of nine features was established based on the previously mentioned methods and was further approved by the algorithm SFFS with JM distance. With this subset, it appeared possible to approach the maximal discriminating power (JM distance of 1,4) in a consistent manner over time.

2.2.2. Calibration of the Model

After having determined a feature subset that enables the discrimination of bottom types encountered in SB across time, classification models were tested in order to find the most accurate model for predicting water in SB when applied to a time series. The following variable were considered in the calibration of the model :

- the nature and number of training areas,
- the bottom types (labels) predicted,
- the features forming the input dataset,
- the number of calibration/validation images from which reflectance have been extracted.

All the tested models were trained with a constant number of trees (100) and an optimized number

of variables used by the model at each node. The following section presents the outcomes of the calibration process.

The first variable to be evaluated was the number and the nature of the areas used for the model training. The four considered situations are the following :

- Situation 1 : all excellent SB (31)
- Situation 2 : only homogenous excellent SB (24)
- Situation 3 : only homogenous excellent SB and other training areas representatives of the following classes (concrete, vegetation and water) (59)
- Situation 4 : only non-SB training areas (104) representatives of the following classes (water, coniferous forest, deciduous forest, cultivated fields, grasslands and urban areas), aggregated to represent pure SB bottom types : concrete, vegetation and water.

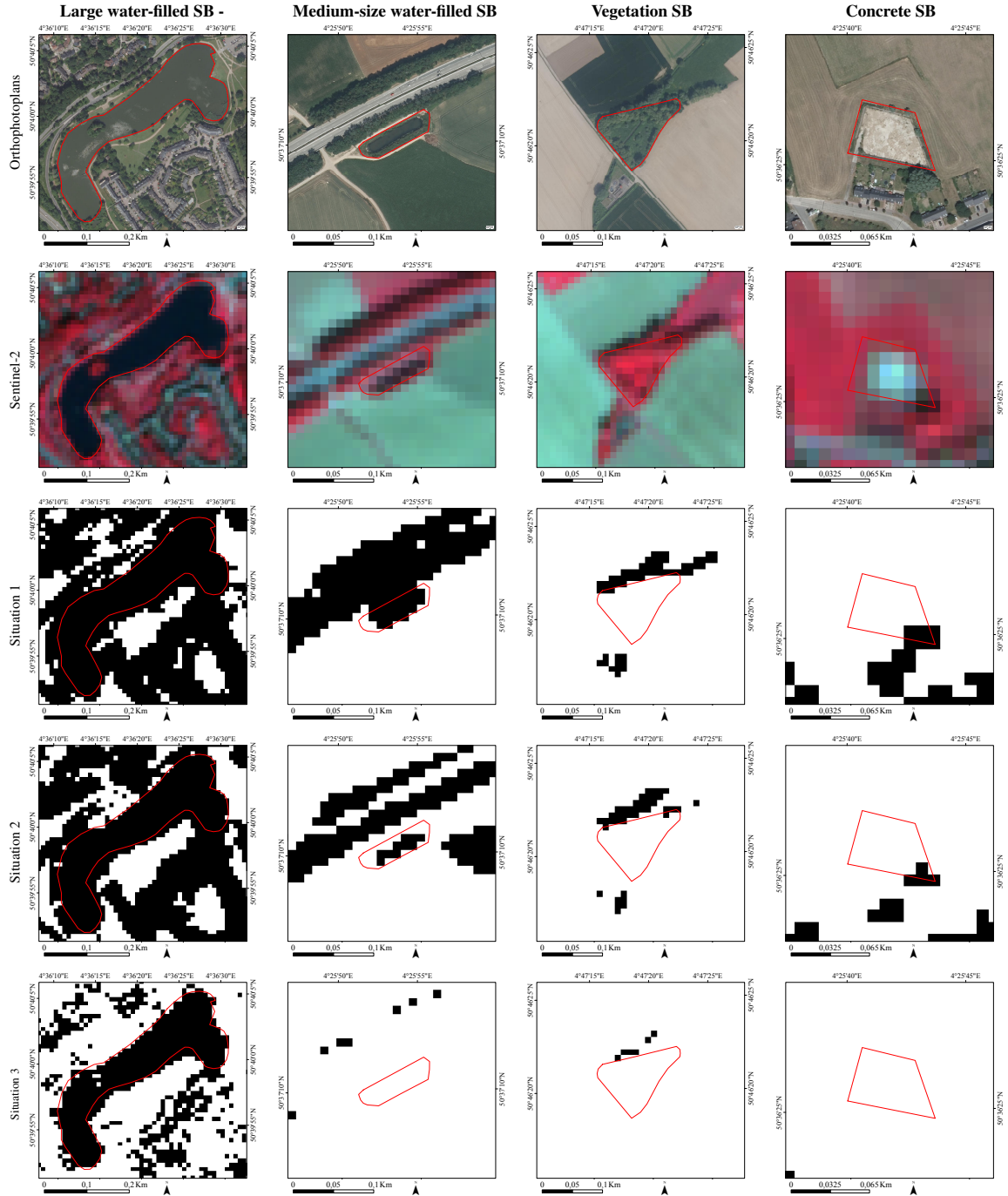
The other variables were kept identical for all the tested models. The models were trained on the reflectance extracted from the September image (18/09/18) for the feature's subset defined at point 2.2.1., in order to predict two categories : "Water – Geomembrane" and "Vegetation – Concrete". The models were further used to classify the September image (18/09/18). This classification product was used to quantitatively assess and spatially visualize the performance of the model.

As a result, the visual comparison of classification products revealed that the reduction of training areas to homogenous SB exclusively (Situation 2) allowed a more accurate representation of the reality in comparison with the situation 1. In fact, overestimations observed in the situation 1 disappeared in the situation 2. However, this reduction of training areas did not solve the misclassification between certain kinds of urban, woodlands and cultivated areas with water, as observed in situation 1. Fortunately, this has only been observed in the surroundings of SB, not inside the SB. The only confusion observed inside the SB in models 1 and 2 has to do with the shadows wrongly classified as water. Reducing the training areas to 24 SB increased the purity of SB for calibration, but at the same time decreased the number of samples used for the training from 1000 to 700 pixels.

To mitigate this problematic limited number of representative samples, supplementary training areas representatives of the bottom types were added. As observed in Figure 20, the addition of such supplementary training areas to homogeneous SB enabled the improvement of the classification product outside of the SB by reducing the confusion between water and urban, woodlands and cultivated areas, as well as inside the SB by mitigating the confusion with the shadows. However, this seemed to negatively impact the accuracy of water detection since several cases of underestimation were highlighted in excellent basins. These underestimations of water in excellent SB suggested that this model 3 could possibly prevent the detection of water in medium basins. Same but amplified conclusion as model 3 has been observed in Situation 4. In fact, training the classification model with only representative training areas of the bottom types and without SB (Situation 4) enabled a considerable reduction of the confusion between water and urban, woodlands and cultivated areas, as well as inside the SB by mitigating the confusion with the shadows but this is balanced with the serious underestimation in SB and even the non-detection of water in smaller excellent basins. The confusion matrices of the four situations are displayed in the annexes (Table D. 1).

These results suggest that confusion of water with the surroundings of SB is unavoidable if the model has to be able to detect water in small SB. Taking this into account, one can state that the second model

trained on homogeneous basins constitutes the best option of the number and nature of training areas, as it provides the most pertinent estimation of water inside the basins. While adopting this model, it should be mentioned that it is sensitive to confusion with shadows inside SB and with urban and woody areas outside the SB's extent. One can also deplore the small number (24) of homogeneous excellent SB in the study zone which has led to the training of the classification model on a limited quantity of pixels (700).



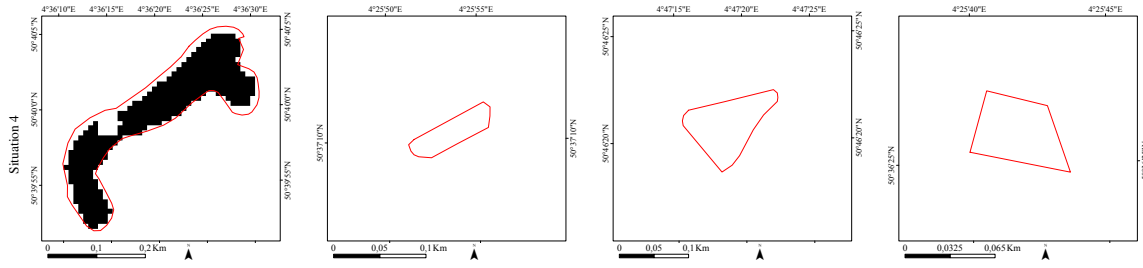


FIGURE 20 – Reduction of confusion outside the SB with an increasing number and homogeneity of the training areas (Situations 1 to 4), as well as inside the SB, an important and slight overestimation of water in situations 1 and 2 combined with a severe underestimation of water in situations 3 and 4 respectively.

The second parameter to be calibrated was the number of bottom types predicted by the model. A choice was made to predict two categories (“Water-Geomembrane” and “Vegetation-Concrete”) instead of trying to discriminate three categories (“Water-Geomembrane”, “Vegetation” and “Concrete”). This choice has been firstly dictated by the fact that the strict objective of this thesis (i.e. the monitoring of water inside SB) only require to discriminate water from other bottom types. Since it is not possible to discriminate water from geomembrane, the best course of action was to discriminate between “Water-Geomembrane” and “Vegetation-Concrete”. Secondly, when considering the 31 excellent SB as training areas, it was observed that classification models trained to predict two categories (“Water-Geomembrane” and “Vegetation-Concrete”) were performing better than models trained to predict three categories. Lastly, even if we wanted to predict 3 categories, the amount of concrete training areas when opting for 24 homogeneous SB was too low to correctly assess the ability of the model to predict three categories. Consequently, the model developed in the rest of the work will be trained to predict two categories because predicting two categories serves the purpose of monitoring the water surface while providing a better accuracy on classification.

Finally, eight models were developed in order to simultaneously assess the variables regarding the number of features and images (from which reflectance measurement are extracted) taken into account in the classification process. Therefore, the models were trained either with the full set of features or with the subset of features determined previously. Reflectance observations may have been extracted from one of the calibration/validation images or from the three simultaneously. Regarding the other variables, the training areas were restricted to homogeneous excellent SB and the predicted categories were set up to : “Water-Geomembrane” and “Concrete-Vegetation”. Figure 21 shows the overall accuracy and F1-score. The precise confusion matrices and related metrics shall be set out in the annexes at Table D. 2.

It has been observed that :

- The overall accuracy of all the tested models exceed 85%.
- For each date, the models trained with the full set of features expressed similar accuracy than the models trained with the subset of features since their 95% confidence interval overlapped.
- The models trained on the 18/09/18 highlighted an increased F1-score for the “Water-Geomembrane” category.

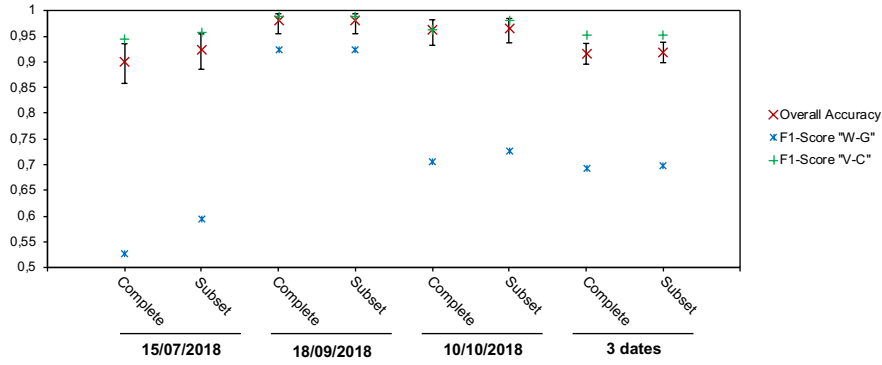


FIGURE 21 – Precision of the Random Forest classification models depending on the reference image (July, September or October image as well as the three images simultaneously) and the number of features (subset or full set) selected for their training.

Since their accuracy are similar, a choice was made to work with the subset of feature as it partially prevents the overfitting of the model on the observations. Among the model tested on the subset of features, overall accuracy metrics indicated that the model trained on the September image (18/09/18) was the most performing. When classifying the three calibration/validation images with the models trained with the September image and the one trained with these three images, one can observe that the confusion matrices and related metrics (Annexes at Table D. 3) indicates a comparable accuracy of classification among models. In contrast, the visual inspection of the three classified calibration/validations images respectively revealed a better representation of the reality when the model used had been trained with the September image than the one trained with the three date images. Consequently, the mono-date approach for model training has been preferred over the multi-date training approach.

2.2.3. Conclusion

To sum up, the inspection of classification product and confusion matrices both indicate that the most effective Random Forest classification model is the one trained with reflectance extracted from the 24 homogeneous SB training areas in the September image with the aim to predict two categories (“Water-Geomembrane” and “Vegetation-Concrete”) with the following features only : Band 2, Band 4, Band 8A, Band 11, NDWI, AWEInsh, AWEIsh, LSWI and s.

2.3. Validation of the Model over Excellent and Medium SB

The following section presents the outcomes of the performance assessment (strengths and weaknesses) of the classification model previously retained (Section 2.2.1.) in predicting the correct bottom types in medium and excellent basins, when using the reflectance extracted from the calibration/validation images. The model performance was assessed by means of a traditional comparison of the predicted bottom type with the corresponding reference for each pixel, as well as with a comparison of the most frequently predicted bottom type and the corresponding reference at the level of SB. In addition to these quantitative approaches, a visual comparison of the initial Sentinel-2 image with its classification was conducted in order to spatially assess the classification validity. These steps were performed for the following categories of SB : excellent homogeneous, excellent heterogeneous and medium.

2.3.1. Homogeneous Excellent SB

When the model is applied to predict the bottom type of homogeneous excellent SB used for the model training on the three reference images (among which 2 of them did not serve the model training), it is noticed that :

- The model predicts the bottom type for each pixel with an overall accuracy of 96% (Table D. 4.a).
- The predicted modal class corresponds to the reference in 93% of the cases (Table D. 4.b).

The visual analysis of classification products demonstrated that all bottom types (water, concrete vegetation) are correctly predicted by the model, as shown in Figure 22. The only drawback to be reported for homogeneous excellent SB is the misclassification of shadows in the “Water-Geomembrane” category. In contrast to what is seen in Figure 22 , in most circumstances, the impact of such misclassification remains marginal on such SB thanks to their size/shape and may even be canceled by the presence of a negative buffer of 10m on SB.

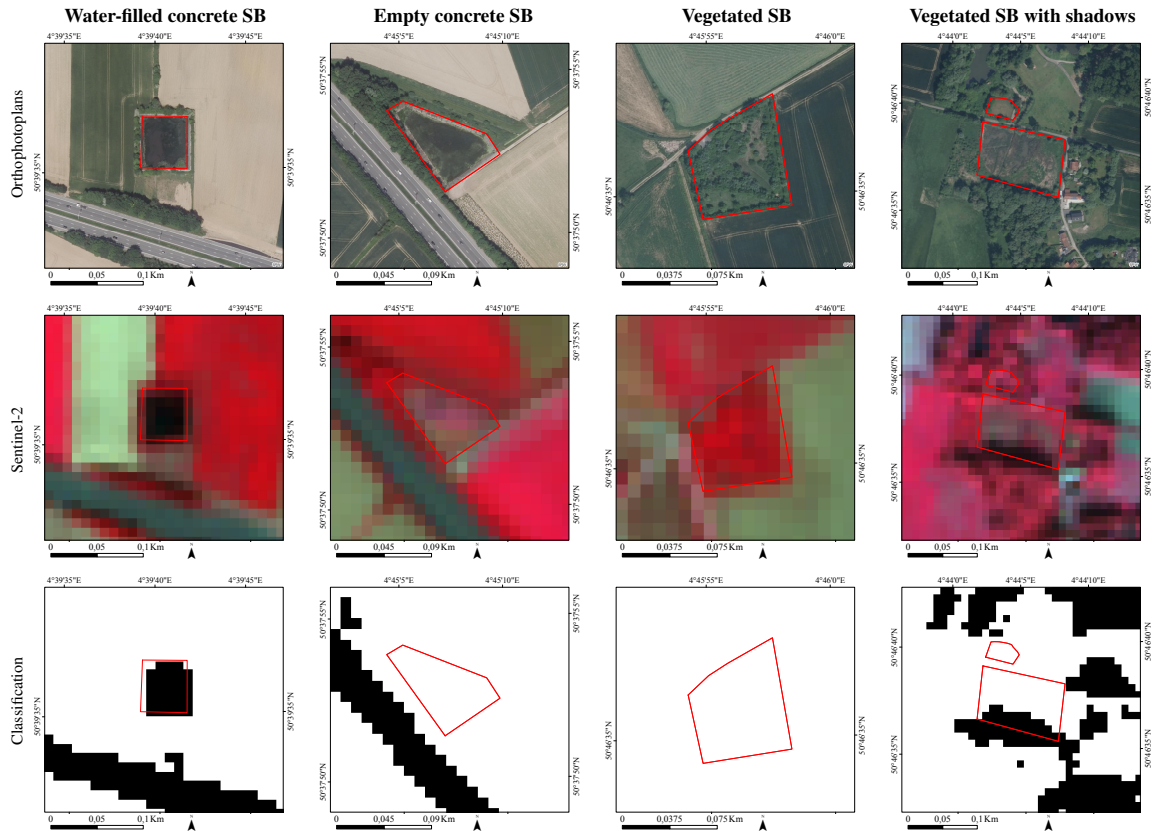


FIGURE 22 – Ability of the model to depict the correct bottom types (water, concrete and vegetation) inside the homogeneous excellent SB, except in presence of shadows. A severe confusion between shadows and water is illustrated in a vegetated homogeneous excellent SB.

2.3.2. Heterogeneous Excellent SB

When the model is applied on the seven excellent SB not used for the model training due to the heterogeneity of their surface, one can quantitatively observe :

- The model predicts the bottom type for each pixel with an overall accuracy of 78% (Table D. 4.c)
- The predicted modal class corresponds to the reference in 62% of the cases (Table D. 4.d)

The quantitative metrics of accuracy suggest that the model is less effective when applied on heterogeneous SB than on homogeneous SB. This statement has been overturned by the visual analysis which indicates that the model reports the heterogeneous parts of SB (Figure 23) with the same degree of accuracy as for homogeneous excellent SB. This further highlights the limited capacity of quantitative indicators in assessing the classification accuracy when there are heterogeneous conditions. This is due to the difference in the level of detail between the predicted and reference values. In fact, a single reference value was determined for each basin and attributed to its pixels without considering their individual spectral signature. In contrast, a predicted value was assigned to each pixel through a classification process based on its individual spectral signature. Normally, the modal class should have tackled this problem by adjusting the level of detail of the predicted values to the one of the reference values. However, the results surprisingly indicated a mismatch between the predicted and reference modal class several times across calibration/validation images for the 4 basins illustrated in Figure 23. However, the illustrations suggest that the model enables a good representation of the reality for all of them. Further investigations demonstrated that these SB counted almost a half of each bottom type. Consequently, the predicted modal class resulted from a tiny difference of frequency among bottom type.

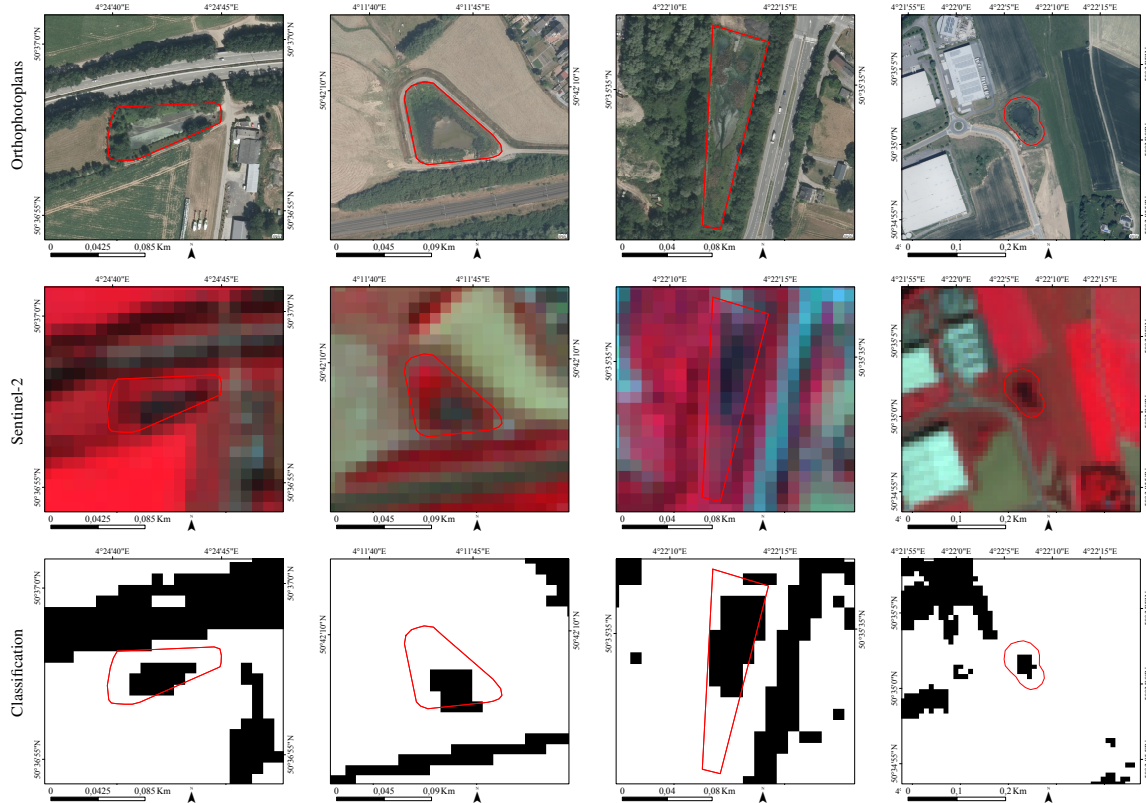


FIGURE 23 – Ability of the model to depict the heterogeneity inside the heterogeneous excellent SB.

2.3.3. Medium SB

Regarding the model performance when applied on medium SB, the model express :

- The model predicts the bottom type for each pixel with an overall accuracy of 72% (Table D. 4.e).
- The predicted modal class corresponds to the reference in 60% of the cases (Table D. 4.f).

The quantitative metrics of accuracy suggest that the model is less powerful when applied on medium SB than on excellent SB. This is explained by the fact that medium basins are harder to identify than excellent ones as they are subject to at least one constraint regarding the visibility of their surface (between 20 and 90% of their surface is visible) and their size/shape (between 1 and 9 pixels remaining after the buffer). Based on these definitions, one can intuitively state that the medium basins category comprises SB of very diverse appearances, ranging from basins with one constraint of visibility or size/shape to basins with two constraints. The modal prediction is even worse than with excellent homogeneous SB since the calculation of the modal class was performed on fewer observations, which makes it relatively unstable.

The visual analysis of the classified time series images revealed outcomes as diverse as their SB appearances. For the ease of understanding, the following section presents the strengths and limits for 1) basins with one constraint of visibility and 2) basins with limiting size/shape. It should also give an indication on situations where basins have both constraints.

Among basins having a restriction regarding the visibility of its surface exclusively, one can state that the classification is all the more informative that permanent obscuring elements do not entail shadows or that the entailed shadows remain marginal with regards to the basins' surface. As an illustration, reeds and low bushes (Figure 24) constitute obscuring elements of the surface but they do not compromise the classification accuracy since they do not entail (or very few) shadows. In contrast, tall trees inside or at the edge of SB introduce a confusion between water and shadows, as illustrated in Figure 24 (last column). This confusion may in some case prevent the assessment of the classification accuracy as it is impossible to distinguish shadows from water.

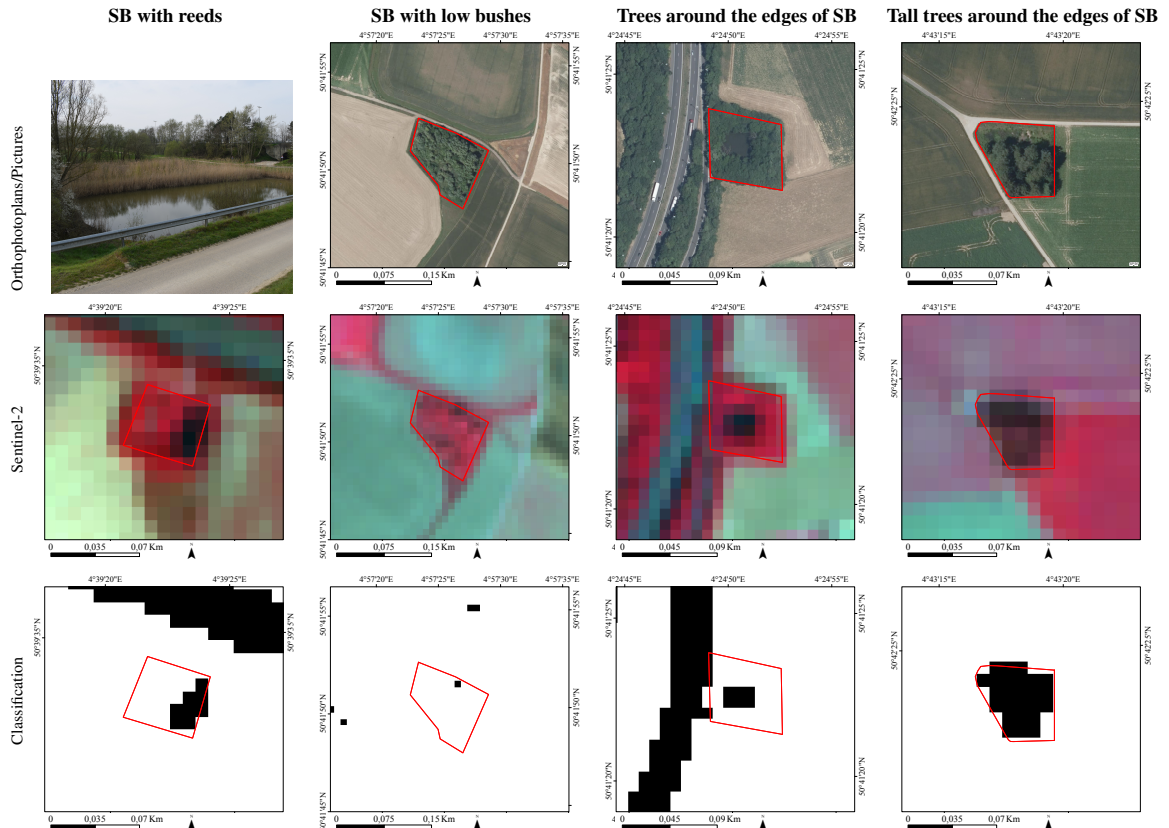


FIGURE 24 – Ability of the model to depict the correct bottom types (water and vegetation) and the heterogeneity inside the medium SB, except in presence of shadows. A severe confusion between shadows and water is illustrated in a vegetated medium SB.

Among basins having a restriction regarding the size/shape of its surface exclusively, one can state that the classification is all the more informative that the remaining pixels are "pure" on a contrasting background. It is also illustrated in Figure 25 that the neighbourhood effects are more severely affecting the SB surface when their size/shape ratio is low.

The statements regarding the major determinant of the informative character of the classification add up when the SB is subject simultaneously to the size/shape and the visibility constraints.

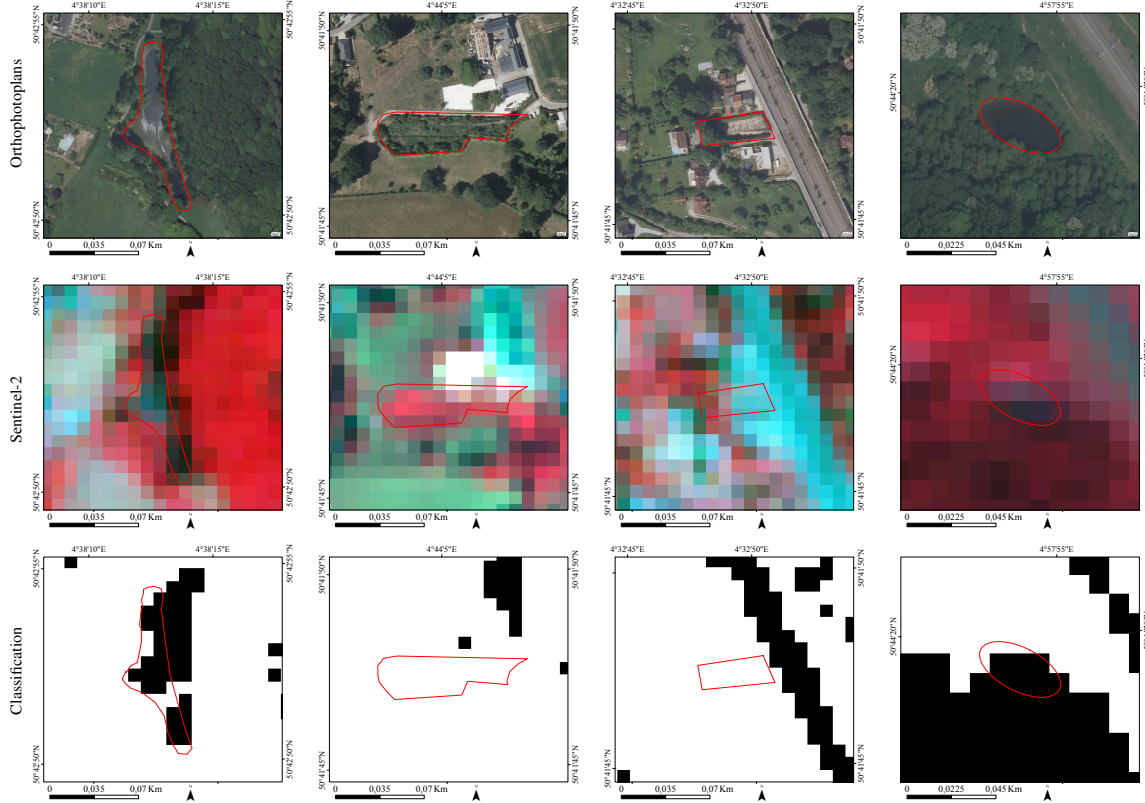


FIGURE 25 – Ability of the model to depict the correct bottom types (water and vegetation) in even small SB, provided that the background is not contaminating.

2.3.4. Conclusion

To conclude, the strength of the model lies in its capability to accurately ($>72\%$) represent bottom types for both excellent and medium SB, even when they are subject to heterogeneity. The performance of the model for medium SB (72%) is reduced in comparison with excellent SB (96% for homogeneous and 78% for heterogeneous⁶) since the model has to deal with additional constraints regarding the size/shape and visibility of the basins when it is applied on medium SB. The only limitation of the model identified when analyzing the calibration/validation classified S2-images is the misclassification of water caused by the presence of shadows which are resulting from the presence of permanent obscuring elements (like vegetation, dams, etc.). This problematic situation appeared to affect more severely medium than excellent SB because of the higher relative surface area of the shadow relatively to the SB surface. Taking this limitation into account, the model demonstrated sufficient performance to serve the purpose of monitoring SB.

6. The accuracy of heterogeneous is reduced (78%) according to the quantitative accuracy metrics only.

3. Monitoring of Water Inside SB

3.1. Application of the Model over the Time Series for both Medium and Excellent SB

In order to monitor SB, the validated classification model was applied on pixel reflectance values extracted from the Sentinel-2 image time series in order to predict the temporal variation of the proportion of pixels classified as water in each SB.

3.1.1. Time Series of Classified Sentinel-2 Images : Strengths & Weaknesses

In the first instance, the visual analysis of the time series of classified Sentinel-2 images demonstrated that the model succeeded in determining the expected⁷ water coverage in excellent SB across the time series, as illustrated in Figures 26 and 27. The same was observed for medium SB.

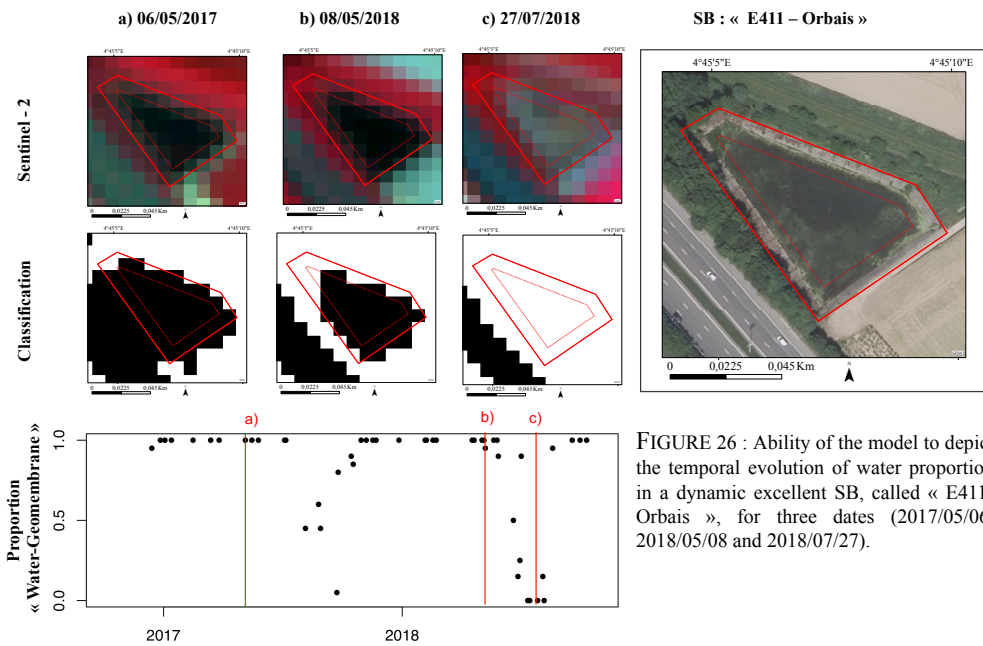


FIGURE 26 : Ability of the model to depict the temporal evolution of water proportion in a dynamic excellent SB, called « E411-Orbais », for three dates (2017/05/06, 2018/05/08 and 2018/07/27).

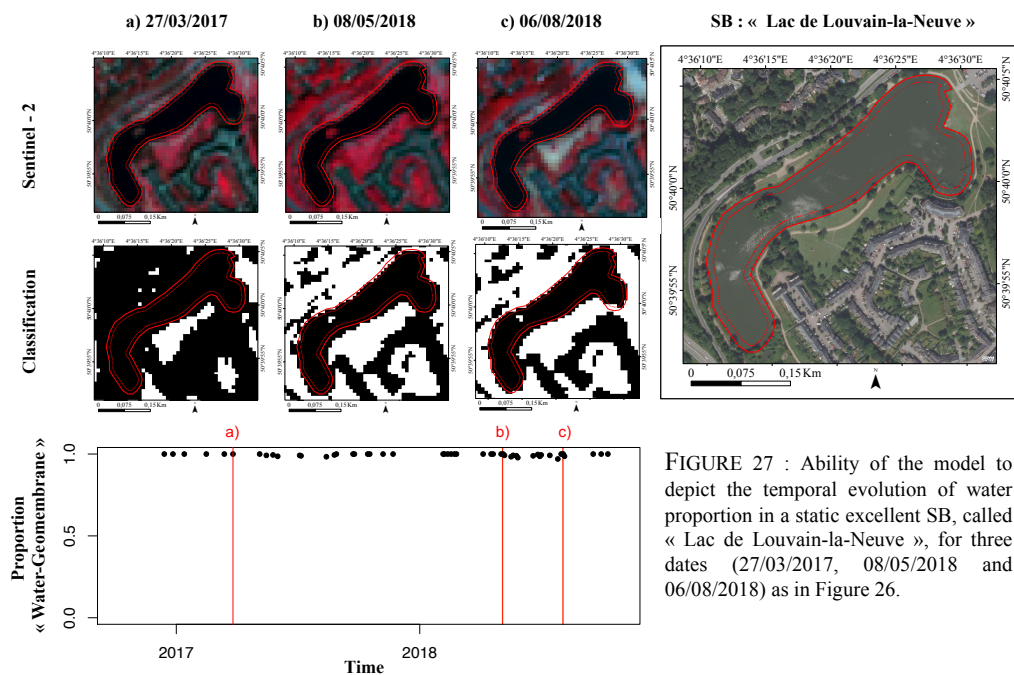


FIGURE 27 : Ability of the model to depict the temporal evolution of water proportion in a static excellent SB, called « Lac de Louvain-la-Neuve », for three dates (27/03/2017, 08/05/2018 and 06/08/2018) as in Figure 26.

7. The water coverage was expected based on the spectral signature of the SB in false-color composite.

In the second instance, the evaluation of this time series of classified Sentinel-2 images revealed additional limitations to the the confusion between the shadows and water identified during the analysis of the classified calibration/validation images. The rest of this section aims to describe them. The quantification of their corresponding errors is comprised at section 3.2..

The first limitation of the classification model occurs when the land cover of Sentinel-2 images to be classified is covered by snow. In such cases, pixels tend to be classified as "water", regardless of the actual land cover. This biases the resulting proportion by overestimating it (Figure 28 first column).

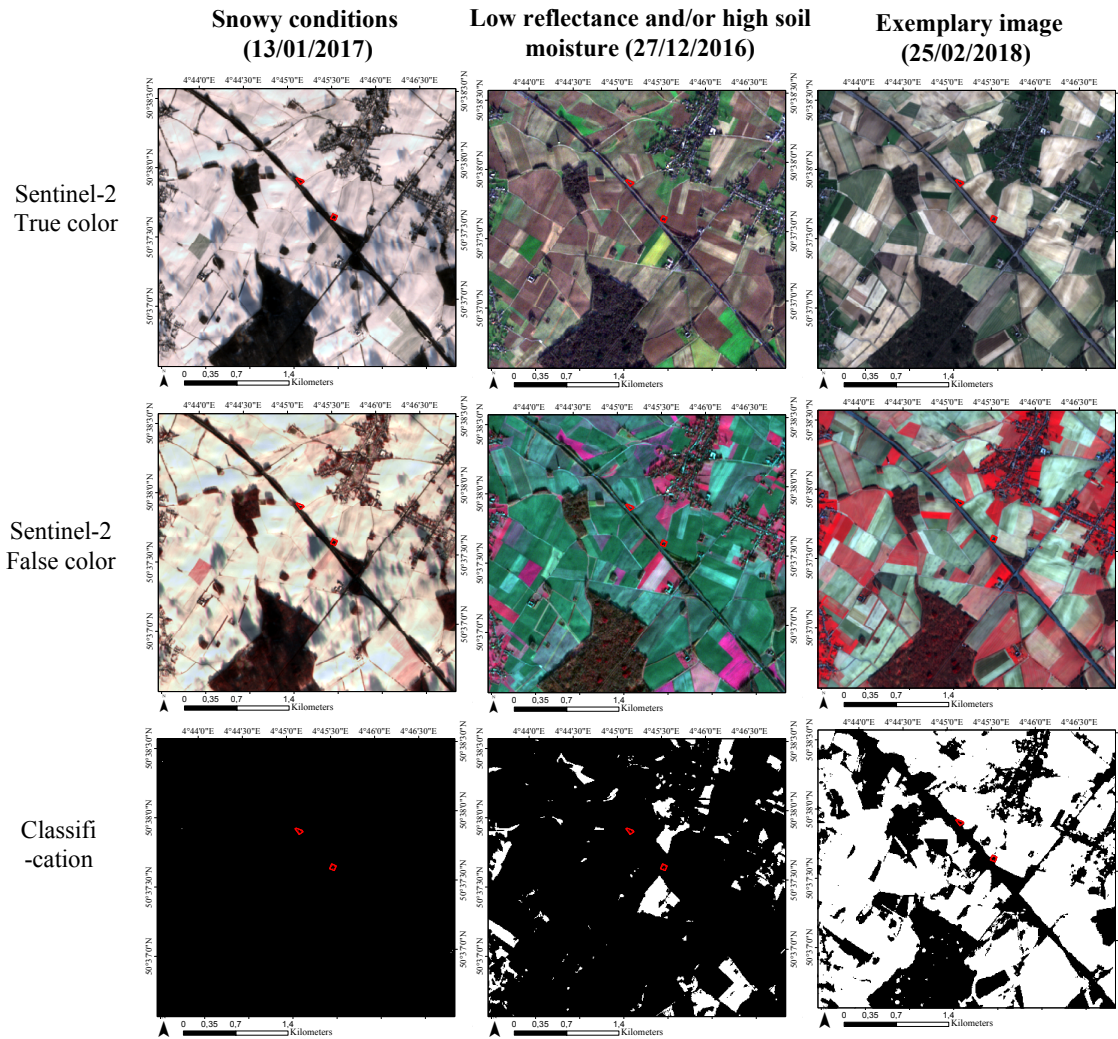


FIGURE 28 – Overestimation of water resulting from the presence of a snow cover and from the lower reflectance and/or higher soil moisture Sentinel-2 images in comparison with an exemplary image acquired during the same type of season.

A second possible source of overestimation of water is induced by the lower reflectance in comparison with an exemplary image at punctual dates during winter months (Figure 28 second and third column). This can be related simultaneously to the higher soil moisture or to the lower sun's course in winter.

A third possible source of error regarding the water proportion estimation has to do with the residual geometric correction. As mentioned previously in Chapter IV.2.2., the latter causes a shift of the landscape in images across time. The shift causes the underestimation of the proportion as illustrated in Figure 29, where the real water surface area remains constant but the shift artificially reduces the portion of water surface within the limits of the polygon.

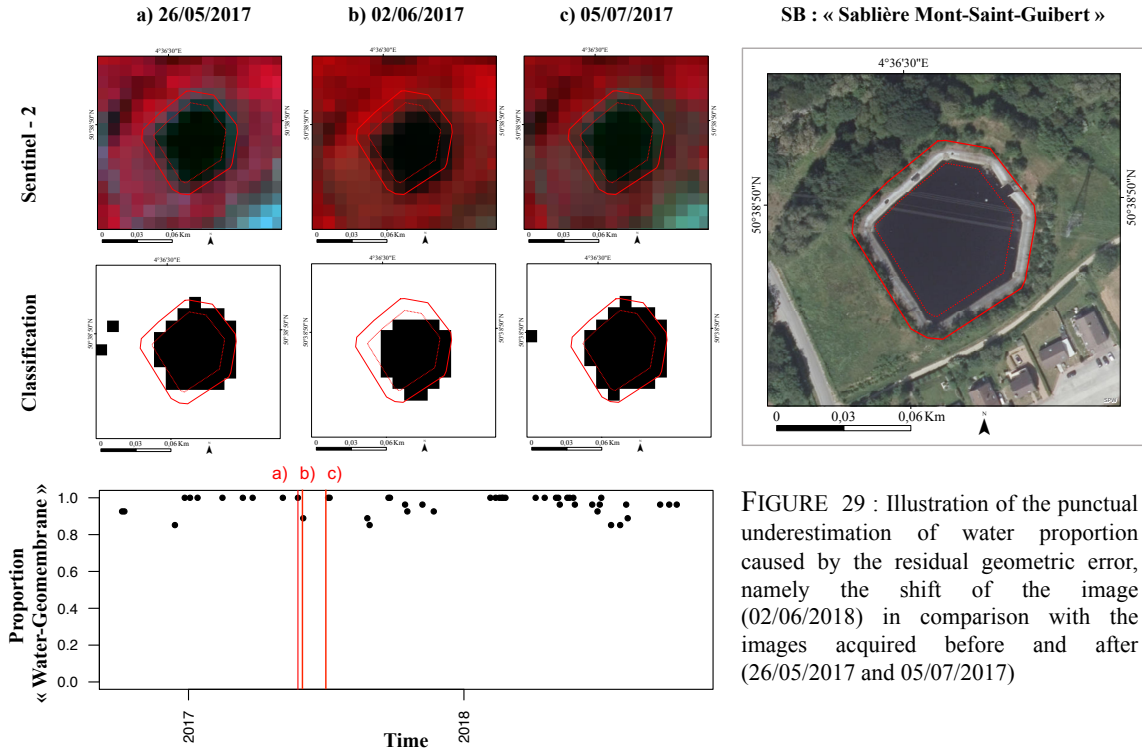


FIGURE 29 : Illustration of the punctual underestimation of water proportion caused by the residual geometric error, namely the shift of the image (02/06/2018) in comparison with the images acquired before and after (26/05/2017 and 05/07/2017)

Fourth, the transient presence of algae on the SB surface acts like an obscuring element which further causes the underestimation of the water proportion. The examination of orthophotoplans and Sentinel-2 images revealed a pertinent illustration of this phenomenon, which has also been observed in situ in the same SB (but not at the same time). As illustrated in Figure 30, the part of SB surface covered with algae is not recognized as water, despite the presence of water underneath. This is resulting in an underestimated water proportion. An example of SB fully covered with algae is displayed in the annexes (Figure E. 2).

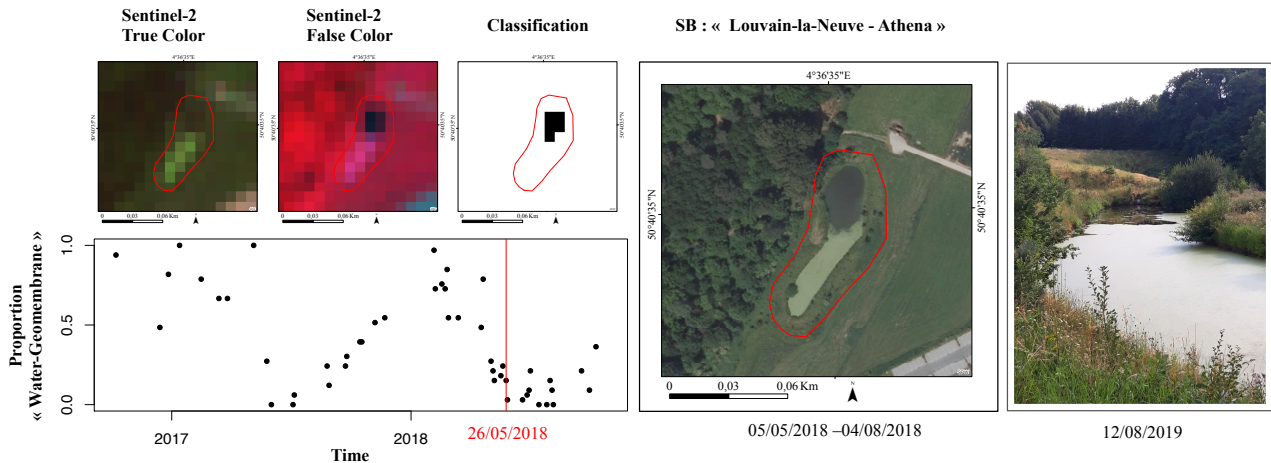


FIGURE 30 – Illustration of the occasional underestimation of water proportion caused by algae bloom.

3.1.2. Reduction of Frequency of Observations by Clouds

The time series used for the monitoring of SB were initially composed of 100 and 155 pre-processed images for the UES and UFS tiles respectively. However, not all these images resulted in a proper and exploitable observation. In fact, the clouds and missing values removal (that took place during the images pre-processing as described in Chapter V) severely reduced the possibilities to monitor the SB surface. The average number of observations after the clouds and missing values removal amounts

50 observations. The distribution of these 50 observations was uniform throughout the period covered by this study, since the clouds are evenly distributed during the year.

3.1.3. Observed Evolutions

Once classified, these 50 cloud-free observations enabled the computation of the temporal evolution of the water proportion in each basin. The obtained temporal evolutions of the water proportion revealed differences in the evolution of SB in Walloon Brabant. The latter can be categorized into 4 groups according to their stability/dynamism and to their elementary state (presence/absence of water) : 1) Static dry, 2) Dynamic dry, 3) Dynamic wet and 4) Static wet.

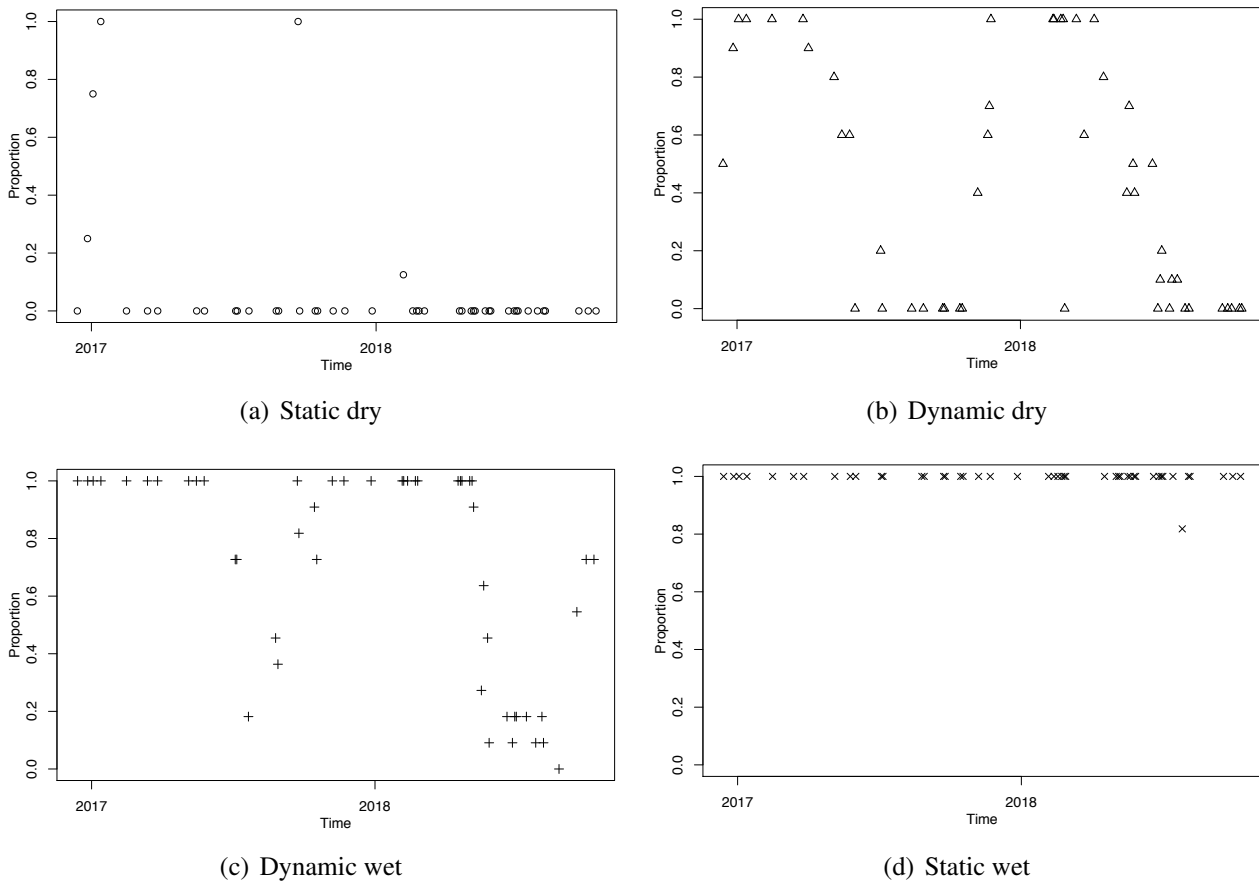


FIGURE 31 – Typical temporal evolution of water proportion observed in each group (Static dry, dynamic dry, dynamic wet and static wet) of the SB.

The comparison between medium and excellent SB evolutions brought to light a more important variability for the medium SB. This may be related to the lower number of pixels contributing to the proportion estimate, as well as, to the reduced performance of the model when applied on more constraining SB.

Nevertheless, for both types of SB, the obtained evolutions provide valuable information regarding the elementary state and the temporal evolution of their water surface.

As a proof, the next figures (Figure 31,32 and 33) summarizes the achievements and limitations of the developed model by showing an obtained classification of water (in black) in space for closely neighbouring SB, as well as their temporal evolution of water.

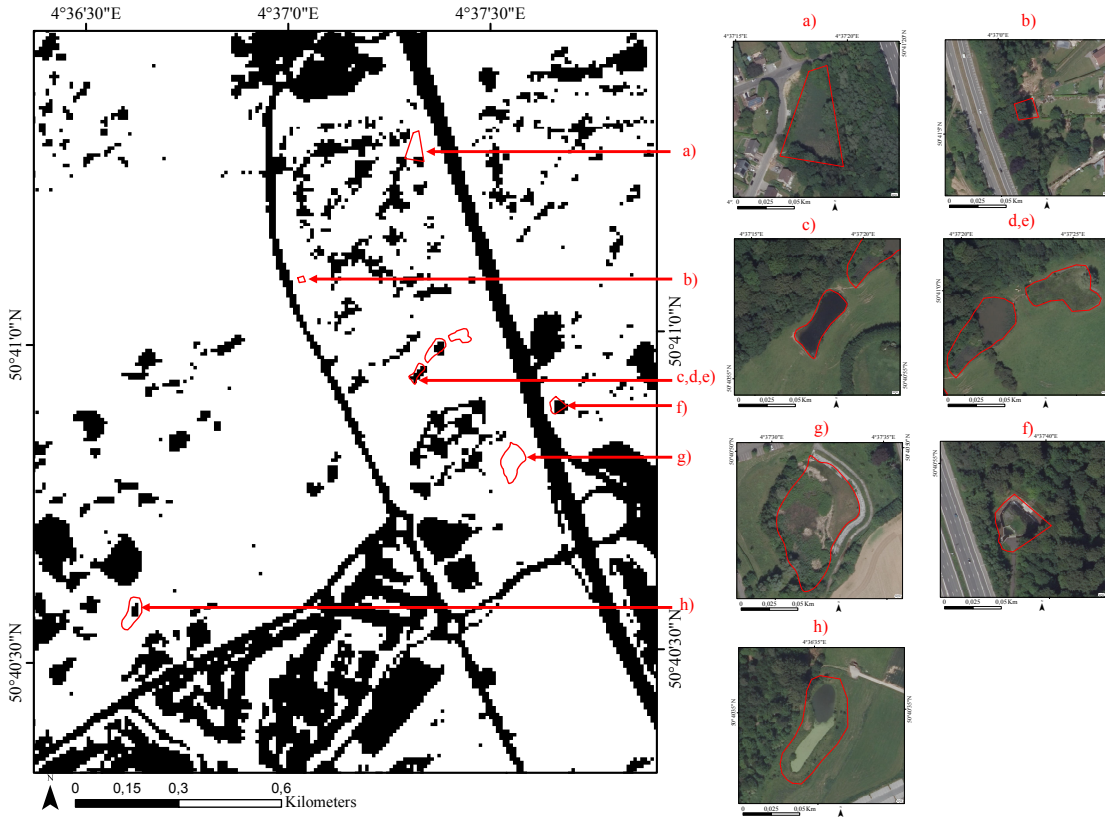


FIGURE 32 – Diversity of SB's state across the classified Sentinel-2 image acquired on July 02,2018 on a part of Louvain-la-Neuve and Chaumont-Gistoux communes. The excellent (a,g) and medium SB (c,d,e,h) are correctly predicted, even though a) and h) suffer from the presence of shadows and algae respectively (more detail in Figure E.2). SB f) illustrates the confusion between water and geomembrane whereas b) demonstrates the difficulty to observe challenging SB on both orthophotoplans and Sentinel-2 images.

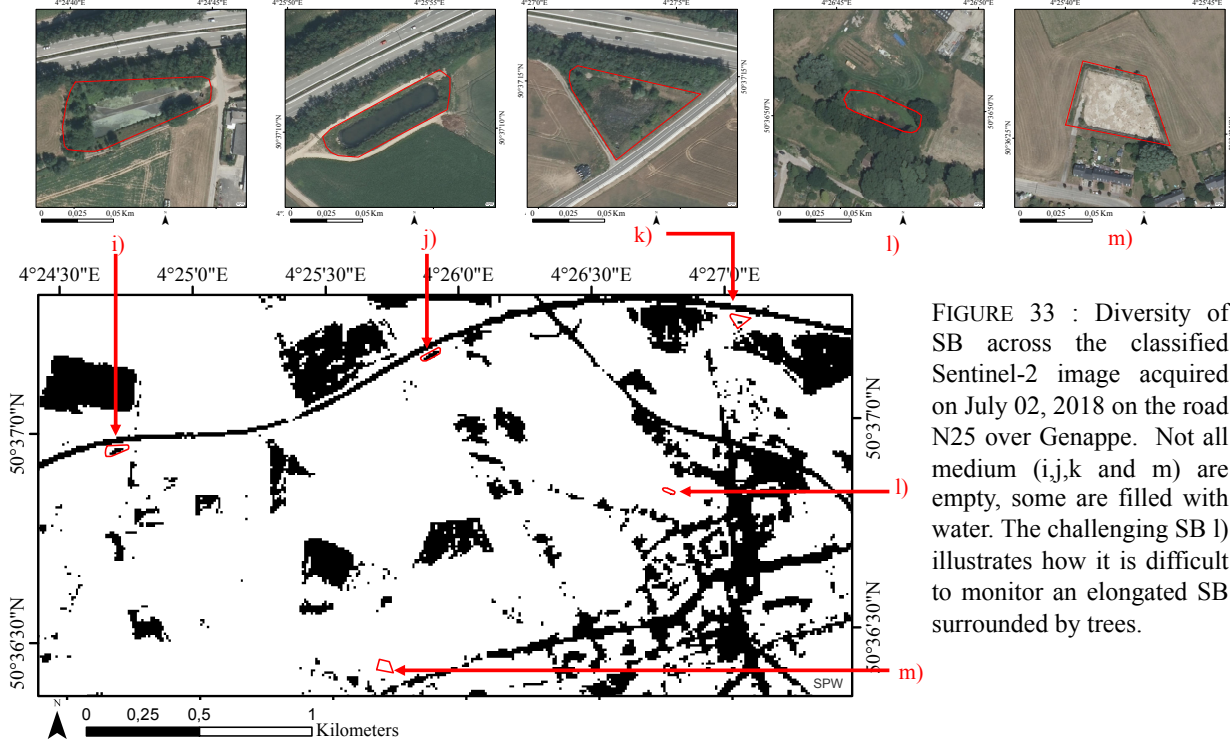


FIGURE 33 : Diversity of SB across the classified Sentinel-2 image acquired on July 02, 2018 on the road N25 over Genappe. Not all medium (i,j,k and m) are empty, some are filled with water. The challenging SB l) illustrates how it is difficult to monitor an elongated SB surrounded by trees.

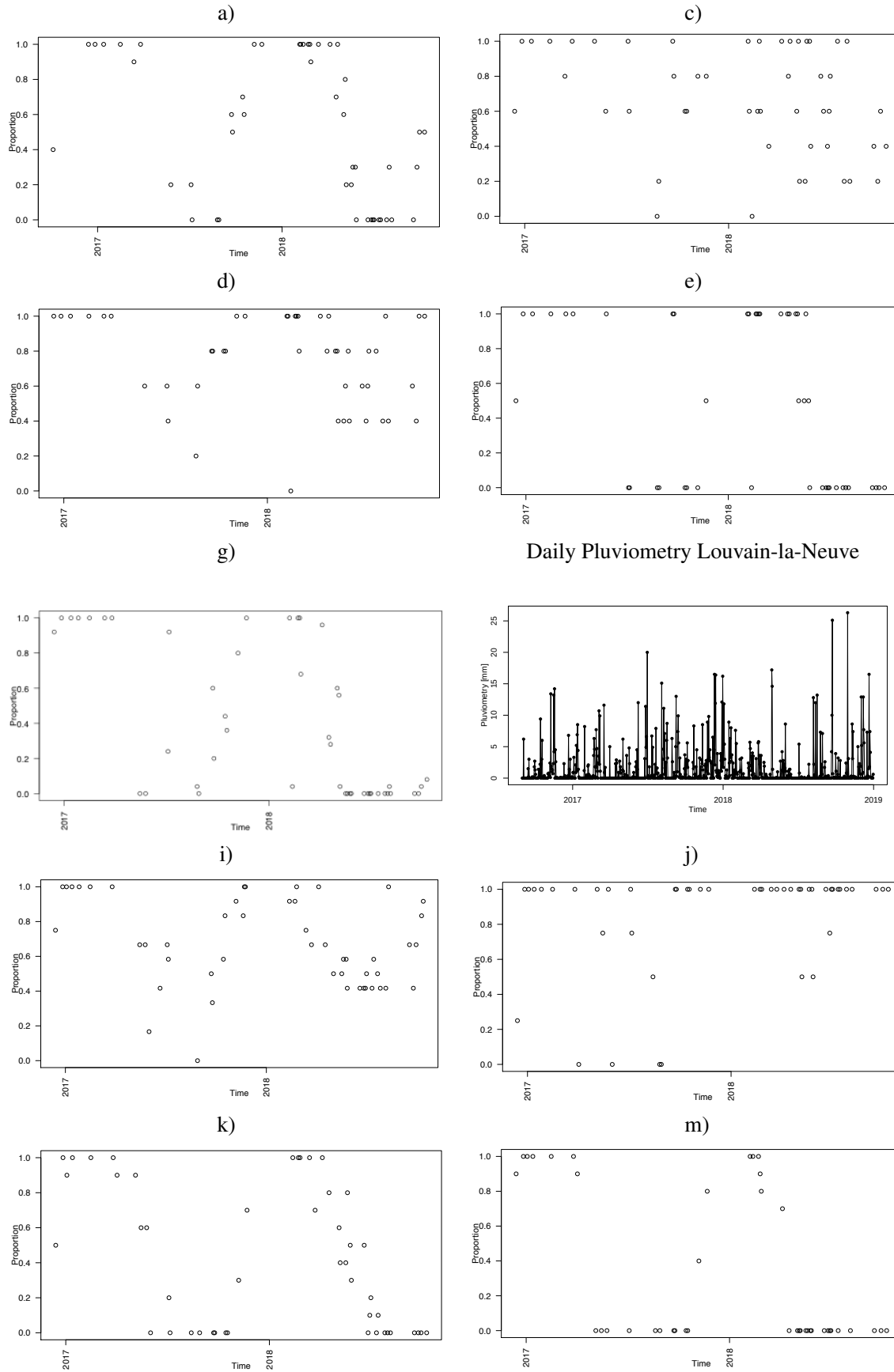


FIGURE 34 – Illustration of the diversity of water temporal evolution encountered inside the SB depicted in Figures 32 and 33 (same numbering). One can see that neighbouring SB like a,c,d,e,g react differently to the pluviometry of Louvain-la-Neuve.

3.2. Validation of the Temporal Evolutions

The validation of the temporal evolutions resulting from the monitoring was performed first by assessing the adequacy between the initial image and its classification. This has been done once on all the 50 cloud-free observations for 10 medium SB and 10 excellent SB (temporal approach) and second time on all basins for 10 images (spatial approach). In addition, a comparison of the water temporal evolutions obtained with Sentinel-2 time series and the one depicted by the aqualim lysimeters has been conducted.

In total, 1660 visual comparisons were performed. The results are shown in Table E. 1 which represents the counting of well and miss-classified images for each SB, and in Tables E. 2 and E. 3 that comprise the number of well and miss-classified SB for each images. The misclassifications were further distinguished according to their causes when possible or were recorded in the comprehensive field “False classification” otherwise.

According to the validation, the model performed with an accuracy of 72% along the image time series of 10 SB and for all SB in 10 images. This step also allowed the quantification of the water proportions error.

- When performing the assessment per images, it was noticed that the error resulting from the presence of snow caused the overestimation of water in all the SB (0% of accuracy) in 1 out of 10 images. When performing the assessment per basin, it was noticed that all the SB were subject to overestimation of their water proportion, caused by the snow cover at most 3 times during the evolution of the basins.
- SB were also overestimated when the images were expressing lower reflectance due to the soil moisture and low sun’s course in winter. This has impacted all SB (<0,02 of accuracy) in 1 out of 10 images and only certain SB (>0,70% of accuracy) in 2 out of 10 images. When performing the assessment per basin, it was noticed that all the SB were subject to overestimation of their water proportion caused by the lower reflectance with a variable frequency ranging from 2 to 8, with an average of 4 times during the evolution of the basins.
- The analysis of the images showed that at most 15 (out of 85) SB were impacted on average 6 (out of 10) time by the shadows of landscape features. When performing the assessment per basin, five SB out of 20 (2 excellent basins and 3 medium basins) were strongly impacted by the presence of shadows. The latter compromised the validity assessment for all the images presenting indication of shadow or water in the SB. When the accuracy assessment was impossible, the SB was considered as misclassified. Therefore, the accuracy of the SB impacted by the shadows is very small.
- Out of the 10 analyzed images, only two SB showed three times signs of algae proliferation during the temporal evolution of its water surface. A third one was identified when analyzing the evolution per basins and 1 of its images were impacted by the presence of algae.

From this first validation method, one can conclude that the model is accurate at 72% in both time and space. The following five sources of error can be identified : the lower reflectance in winter images, the presence of snow, the shadows of landscape features, the residual geometrical error and the proliferation of algae.

The second validation method has been performed on the two following SB out of the 15 SB equipped with lysimeters because three were non-monitorable with Sentinel-2 (challenging detectability) and the other were not delineated as they were considered as temporary immersion area in the provincial platform of flood risk management.

As a result, the water-level evolutions confirmed the validity of the obtained temporal evolutions, as they follow the same trend approximately. The water-level time series only provide a better representation than Sentinel-2 time series for the short-duration rising water levels. This is due to the higher temporal resolution of lysimeters' measurement. Nevertheless, the proportion of water predicted with the Sentinel-2 monitoring tend to one when the water-levels measured inside the SB exceed the lysimeter's reference height (Figure 35.a). In addition, the predicted proportion of water remains null when the water-levels measured inside the SB equals to the lysimeter's reference height (Figure 35.a b).

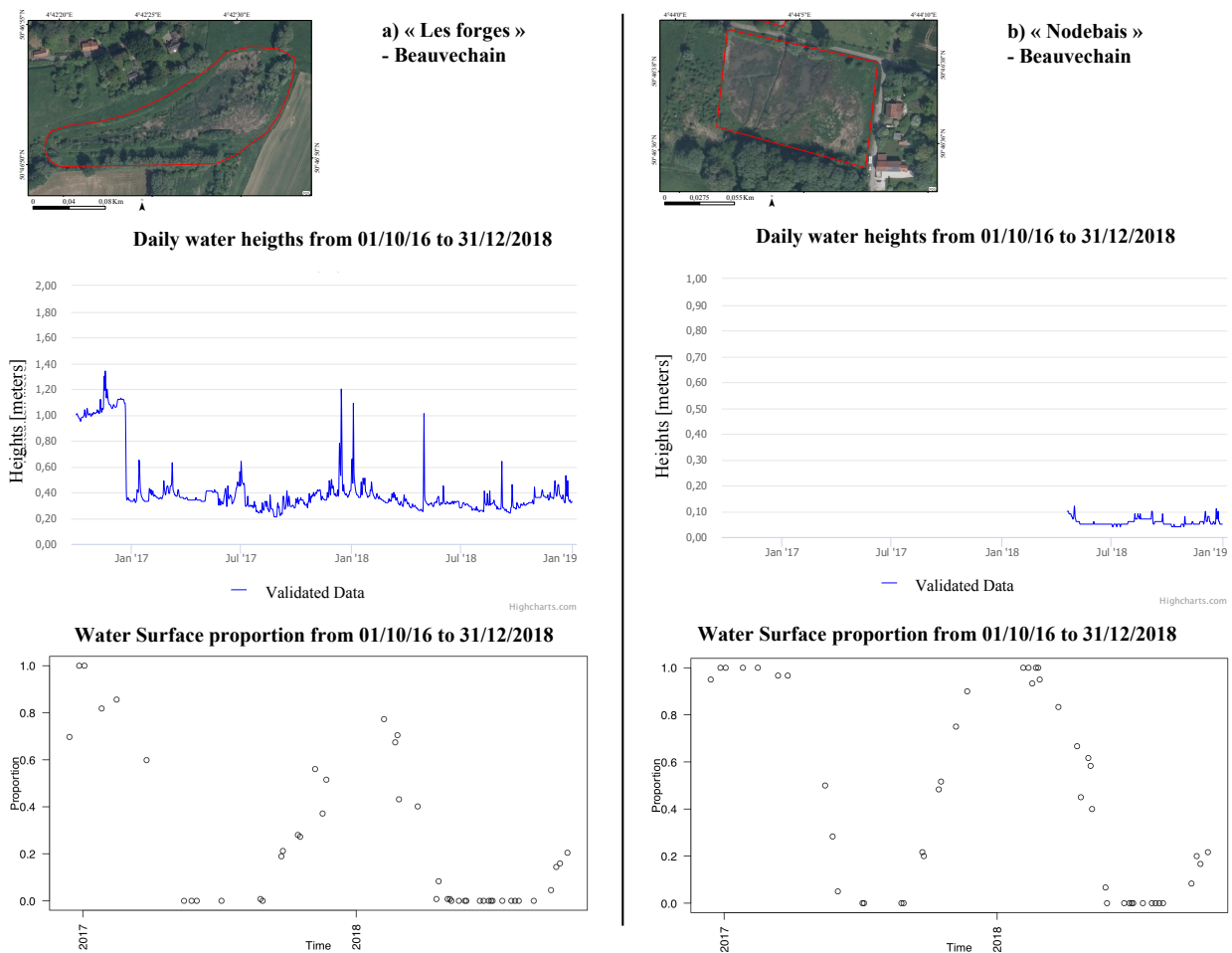


FIGURE 35 – Validation of the obtained temporal evolution with daily water heights inside the SB (Source : Aqualim).

This second validation method confirms that the proposed method of monitoring SB with Sentinel-2 time series performed well.

Chapter VIII. Discussion

This research assessed the feasibility of monitoring SB by means of remote sensing data with the overall objective of deriving valuable information about their hydraulic behavior. The study has been conducted on the basis of Sentinel-2 images with a focus on the geographical area of the province of Walloon Brabant.

To avoid too complex reasoning mainly due to the numerous interconnected variables and to facilitate the reader's understanding, the discussion hereafter will simultaneously address the obtained key findings respectively with regard to the SB detectability and monitoring, and the chosen methodology. This discussion also encompasses the possible prospects for future research.

An in-depth review of the literature on the topic did not reveal any similar studies related to the monitoring of SB. In comparison to other water-related research, the main encountered challenges were the diversity of the SB water characteristics, the poor visibility of SB due to obscuring landscape elements, the diversity in SB's appearance as well as their resemblance with other landscape features. The fact that SB are rather small combined with quite a spread spatial spacing over the study area constitute an additional challenge to be managed in this study.

1. Discussion of SB Detectability

Based on the already acquired experience in that matter and in light of the aforementioned SB characteristics, it could from the beginning, be expected that the Sentinel-2 images might, in terms of spatial and spectral resolution, not with certitude allow the identification of some basins.

Consequently, a methodological step has been implemented to qualify the easiness to observe SB based on Sentinel-2 images. This approach aimed at assigning each SB one of the three chosen levels of detectability (challenging, medium and excellent) by the photo-interpretation of two variables : the size/shape of the SB and the visibility of its surface.

In order to differentiate the levels of detectability related to the size/shape parameter, two threshold values have been settled with regard to the number of pixels remaining after the application of a negative buffer of 10 m : the lowest at the value of one and the other equal to nine. To discriminate the levels of detectability related to the visibility parameter, limit values of 20% and 90% were chosen as threshold percentages of the visible surface in basins.

Once this methodology has been applied to the study case of Walloon Brabant, it was found that the majority of SB (85 out of 140) were potentially observable with Sentinel-2 images since 31 one of them were completely free of constraints and the 54 others were subject to minor restrictions with regard to the size/shape or visibility parameter. In contrast, the detectability of the 55 other SB was highly compromised by either the obtained results regarding size/shape of the SB (lower than 1 pixel remaining after the buffer) or the ones related to the visibility of its surface (lower than 20% of the surface was visible).

On that basis, one can therefore deduct from the not readily observable SB's characteristics that when the resulting buffered polygon encompassing a completely visible surface crosses at least one pixel (i.e. 0,01 ha), the concerned SB presents a sufficient size/shape to be detected by Sentinel-2 images. In the case the size is not a limiting factor, the requirement to be detected by Sentinel-2 images in

terms of visibility is to have at least 20% of its surface visible.

The requirement of size/shape for the identification in SB may be confronted to comparable cases in other studies that present similarities regarding the sensors type as well as the nature, size and shape of the target object.

In that sense, the assessment of Sentinel-2's potential for sub-pixel Landscape feature detection provided by Radoux *et al.* (2016) is very convenient. Radoux *et al.* (2016) have calculated the minimum width required to accurately detect specific foreground objects from their background using the best index or individual band for Sentinel-2, SPOT and Landsat sensors. For the purpose of this study, the minimum widths required to accurately detect foreground objects such as Road, Grassland and Water with Sentinel-2 images, are of particular interest since they can be related to the following SB bottom type : Concrete, Vegetation and Water. Their results demonstrated that the minimum detection width of a square road centered in the pixel ranges from 8 to 18,5 m depending on the considered background with 10 m bands and can increase up to 34 m with 20 m bands of Sentinel-2. When considering square water bodies are centered in the pixel surrounded by grassy background, the minimum detection width is 12,5m and increases up to 19,5 m when using 20 m bands (d'Andrimont, 2017). They also conclude that a square grassland or pasture can be differentiated from the bare soil when the square exceeds 8,5 m for the 10 m Sentinel-2 bands and 14 m for the 20 m Sentinel-2 bands. Their paper has also outlined that it is far more difficult to detect isolated pixels than it is the case with grouped pixels, as their PSF cannot reinforce each other when isolated.

The level of comparison of the outcomes obtained in the present study with the one of Radoux *et al.* (2016) is limited by the fact that their analysis was performed on mono-temporal image whereas the present approach relies on multi-temporal images. For this reason and because of the spatial dispersion of the SB in the landscape, the application of a buffer appeared to be necessary in order to mitigate both the contamination of the signal from the neighborhood and the shift caused by the residual geometric error. Furthermore, Radoux *et al.* (2016) focus on square objects that do not exactly represent the diversity of shapes shown by the tested objects for the present study.

When comparing the results obtained in the present study to the one obtained in Radoux *et al.* (2016), while taking into account the differences in the respective chosen study approach, one can state that both agree on the fact that to be qualified as identifiable, any SB needs to cover more than one Sentinel-2 pixel. Though, it has to be noted that the exact required minimum width to be able to detect any SB or waterbodies effectively depends on the nature and shape of the object as well as on the background to which the SB should be distinguished from. These statements further confirm the well-posedness of the settled thresholds for the parameter of size/shape. It also strengthened the confidence in the fact that the Sentinel-2 MSI sensor seems to be highly adequate for this study since it enables the detection of the majority of SB with its 10m pixels.

In addition, this detectability study-step was useful to select the readily observable SB (excellent and medium) as a suitable base to conduct the attempt of monitoring, as well as to further consider possibilities for the identification of challenging SB. Detecting challenging SB could be done with a combination of techniques to address the defaults of Sentinel-2 images. When the size is a limiting factor, one can recommend performing the assessment with higher spatial resolution sensors ; whereas when the visibility is limiting, one can suggest opting for radar sensors, better able to penetrate through the dense canopy than an optical signal. When both size/shape and visibility are limiting,

physical observation in the field would still be more adequate.

2. Discussion of the Model Calibration and Validation

The second achievement of this study is the development of a classification model that has demonstrated a satisfactory level of efficiency to derive the temporal evolution of the water surface proportion of within the excellent and medium basins of Walloon Brabant when applied on calibration/validation images.

The calibration of the model showed that the most effective classification model was trained on reflectance extracted from parts of the September 18, 2018 image covered by the homogeneous “excellent” SB, in order to predict two categories : “Water-Geomembrane” and “Vegetation-Concrete”. When applied on the calibration/validation Sentinel-2 images for its validation, the classification model was successful in depicting the reality even with heterogeneous areas for both excellent (78% for heterogeneous and 96% for homogeneous) and medium SB (72%).

One could discuss the fact that the geomembrane and water bottom types have been fused together in one category during the establishment of the classification model. This is because the geomembrane bottom type was found to be impossible to discriminate from the water on the basis of available spectral information provided by Sentinel-2. No mention of this problem has been found in the literature. The distinction between the geomembrane and water is often studied to detect underground leaks of water with radar or electrical methods, but not with an optical or IR signal above the ground surface. In any cases, the consequences of this merger one the monitoring remained minor as the fusion compromised the monitoring of only two SB out of 85. The other SB were not impacted by this operation since water is the only bottom type that is susceptible to be added across time.

Another point subject to discussion, is the fact that, during the calibration of the model, the classification model trained with the September image performed better than the model trained on the 3 dates images, when applied to classify the September image. This outcome is inconsistent with what is generally observed and performed in remote sensing. In fact, it is generally advised to train the model on multiple dates in order to improve robustness of the model across time series. It is likely to be specific to this particular study.

Further on, one can discuss the reliability of the quantitative accuracy obtained. For medium, both visual and quantitative assessment pointed out that the performance of the model in depicting water surface in medium was slightly reduced (72%) in comparison with the excellent SB (78-96%). This has to be caused by the additional constraint on the observation of medium SB in comparison with excellent one. However, when applied on heterogeneous SB, the quantitative metrics of accuracy suggest that the model is less effective (78%) than on homogeneous SB (96%). This statement contrasts with the outcome of the visual analysis which indicates that the model reports the heterogeneous parts of SB with the same degree of accuracy as for homogeneous excellent SB. This suggested that F1-scores and the overall accuracy cannot be entirely relied on to evaluate a classification but need to be complemented by a visual analysis.

3. Discussion of SB Monitoring

The third achievement of this study is the application of the validated classification model on pixel reflectance values extracted from the Sentinel-2 time series in order to predict the temporal variation

of the proportion of pixels classified as water in each SB.

The validation of the temporal evolution of water revealed that the classification model was accurate at 72% in both time and space. The analysis of the classified Sentinel-2 images time series revealed 5 sources of error : the lower reflectance in winter images, the presence of snow, the shadows of landscape features, the residual geometrical error and the proliferation of algae. Note that the lower reflectance observed in winter images was assumed to result from the high soil moisture caused by heavy rain.

The discrimination between water surfaces (either flooded, or permanent water bodies) from the artifacts caused by heavy precipitation and wet snow is acknowledged to be a very difficult task because of their concomitant presence during severe weather events and their similar microwave radar and optical signature. Pulvirenti *et al.* (2014) proposed a promising and automatic method that exploit X-band SAR images, DEM, a land cover map and an optical image to detect the dark objects and classify them following an object-based approach.

Shadow is also frequently hampering the accuracy of water extraction and change detection (Shah-tahmassebi *et al.* 2013). Among the proposed solutions, one can mention the improvement of shadow detection algorithms (Huang *et al.*, 2010; Zhu and Woodcock, 2015). There is also a possibility to estimate the geometric shadow cast of objects based on the surface elevation of the object, as well as the sun and azimuth angles (Radoux, 2010). However, dark surface confusion along with shadows estimations for both clouds and features still remains a challenge for water remote sensing.

The residual geometrical error's limitation arises when dealing with multi-temporal assessment and is resulting from the absence of global reference image (GRI) at 10 m resolution. Fortunately, this issue is about to be tackled by the ongoing establishment of the desired GRI.

The proliferation of algae causes an underestimation of the water proportion as the model classify the algae bloom as vegetation. Zhu *et al.* (2017) have proposed a method for distinguishing algae bloom from floating leaf vegetation. The latter could potentially serve the discrimination between algae bloom and vegetated SB bottom.

One can conclude that these limitations are frequently observed when monitoring inland water surfaces subject to severe weather events at 10 meters resolution. Fortunately, solution exists for each of them but their implementation requires more means and effort.

Despite these hardly avoidable limitations of the classification model, the obtained evolution provides valuable information regarding the elementary state of SB and the temporal presence of their water surface.

Finally, the opportunities to derive the proportion of the water surface have been strongly reduced by clouds, but the remaining 50 observations spread between end of 2016, 2017 and 2018 provided competitive frequency in comparison with the one achieved with the current visual inspection technique or with other possible techniques such as drones.

4. Discussion of the Method

After having discussed the results, one can discuss the limitations of the approach itself. The method demonstrates significant effectiveness in depicting the hydrological functioning of SB, even though it was constrained by the following elements :

- The available number of SB to conduct this study has been dictated by the number of SB registered in the provincial platform of flood risk management of Walloon Brabant, as well as by the ability of Sentinel-2 images to allow their identification. These first two elements have in a way limited the available number of SB for the training of the classification model.
- The qualitative character (presence of water) instead of quantitative (depth of water) character of the method has reduced the possibilities of interpretation of the hydrological functioning of SB.
- The challenging accessibility of validation data has decreased the possibility to perform a completely independent validation.

5. Perspectives

Regarding the prospects for follow-up research, two options could be considered.

The first one is to pursue the development of this method by attempting to mitigate the aforementioned limitations. The following areas have been identified for further exploration :

- Improve the level of completeness of the provincial dataset by compiling different sources of already available information (PASH, etc.) or by launching a detection experiment based on the information derived from the present research.
- Use combined techniques (higher resolution sensors, visual inspection, etc.) to increase the number of readily identifiable SB.
- Investigate the synergistic use of Sentinel-1 radar and Sentinel-2 optical satellites to discriminate geomembrane from water bottom types, to prevent the artifacts caused by heavy precipitation and wet snow and to increase the possibility of observation by sensing through clouds.
- Intensify effort in the enhancement of shadow-screening algorithms.
- Redesign the methodology by incorporating the elevation aspect in order to draw the depth of water.
- Increase the accessibility of water-level data that are currently held by the public administrations.

The second approach is to improve already existing monitoring methods that were previously considered as not sufficiently effective. The proposals are as follows :

- One could place a lysimeter in each basin and perform visual inspection to assess the state of the basins (coverage in vegetation).
- One could place a camera in each basin and perform image analysis to assess the state of the basins (coverage in vegetation) while ensuring their operability.
- Another solution would be to rely on the inhabitants, foresters and farmers frequently present in the surroundings of the SB to perform regular assessment of the SB against payment.

Another last option can be based on the combination of the two above-mentioned approaches that are not mutually exclusive.

Conclusion

In light of today's enhanced attention for environmental issues, this study offers its contribution to the search for an overall improvement of water management. More specifically, by exploring remote sensing capabilities, it aims to develop innovative methods for acquiring information about the location and functioning of stormwater basins (SB) at regional scale. Therefore, it assesses the feasibility of monitoring hydrological functioning of SB using Sentinel-2 image time series (end 2016, 2017, 2018) while focussing the research on the province of Walloon Brabant as a study case.

Quite early in the study work, it appears from a methodological perspective necessary to foresee a step-by-step approach relying on three more specific and sequential objectives. The first one consists in preparing the SB dataset and characterizing ex-ante the potential detectability of the SB based on the adequacy between the SB's characteristics and the Sentinel-2 images limitations. The second aims at developing a classification model that exploits the information provided by three Sentinel-2 calibration/validation images in order to detect the presence of water inside the observable SB. The third and last one investigates the potential of Sentinel-2 image time series along with the developed classification model in order to depict the temporal evolution of water inside observable SB.

The rigorous application of the chosen methodology has then provided enough results to conduct the subsequent planned analysis, which, at the end, demonstrates with sufficient confidence that both overall and specific objectives have been fulfilled.

In order to get the expected overall result, the work starts with providing through the first chapters, the necessary background and building blocks as well as the specific framing for the study. On that basis, different experiments consisting in images interpretation, data computing, calculation and modelling work allowed to realize the initially planned assessments and to discuss them. Chapters five to eight respectively describe and further detail those essential and specific steps.

The first one has led to the formation of the SB dataset by delineating SB and the ex-ante characterization of their detectability according to the adequacy between its characteristics and the limitations of Sentinel-2 images. It was found that a majority (140/156) of SB, initially recorded in the provincial platform of Walloon Brabant, were observable and delineated by means of orthophotoplans. Furthermore, the assessment of SB detectability suggested that a majority (85/140) of SB was potentially observable from Sentinel-2 images, since 31 one of them were completely free of constraints and the 54 others were subject to minor restrictions with regard to the size/shape and/or visibility parameter. The detectability of the 55 other basins was highly compromised by either the obtained results regarding size/shape of the basins (lower than 1 pixel remaining after the buffer) or the ones related to the visibility of its surface (lower than 20% of the surface was visible). These results have further strengthened the confidence in the fact that the Sentinel-2 MSI sensor seems to be highly adequate for this study. This part of the work was also useful to select the observable SB (excellent and medium) as suitable base to conduct the attempt of monitoring, as well as to further consider new possibilities for the identification of challenging SB (higher resolution sensor, radar or visual inspections).

The second part succeeded in developing a Random Forest classification model that exploit the spectral information provided by three Sentinel-2 calibration/validation images in order to detect the presence of water inside the observable SB. The performance assessment of the developed model revealed that it was successful in depicting the reality for both excellent and medium SB, even when

there are heterogeneous areas. The accuracy of the classification model ranges between 78% and 96% for excellent SB and slightly reduced (72%) when applied to medium SB which are less easily identifiable.

The third and last part confirmed the feasibility to derive the temporal evolution of water by applying the model on the time series of Sentinel-2 images. In fact, when applied on the time series, the developed model was able to represent the temporal evolution of water with an accuracy of 72%. The 28% of water proportion misclassification can be attributed to the artifacts caused by heavy precipitation and wet snow, the shadowing effects, the residual geometrical error and the algae bloom. The discussion brought to light that these limitations were frequently encountered limitations in complex inland water remote sensing at 10-m resolution.

Despite these limitations, the obtained evolution provides valuable information regarding the elementary state of SB and the temporal presence of their water surface. The prospects suggested that even finer evolution could be obtained by reviewing the methodology in order to incorporate the elevation aspect to draw the depth of water.

The opportunities to derive the proportion of the water surface have been strongly reduced by clouds, but the remaining 50 observations spread between end of 2016, 2017 and 2018 provided a competitive observation frequency in comparison with the one achieved with the current visual inspection technique or with other possible techniques such as drones.

In conclusion, this study demonstrated that Sentinel-2 image time series provides a cost-free means for acquiring information about the elementary state and the temporal evolution of water presence in the observable SB over the province of Walloon Brabant. Furthermore, the implemented methodology allows to determine a priori whether Sentinel-2 would be able to observe the SB and thus monitor its water surface.

Finally, the developed methodology is replicable to any other region provided that the sensor's resolutions are in sufficient adequacy with the SB characteristics (size/shape and visibility) to enable their identification. The average coverage of cloud and snow of the acquired images, as well as the presence of landscape feature shadow should be taken into account when opting for this method.

In the future, all the presented limitations should be addressed. Future research should focus on integrating Sentinel-1 images to the proposed approach. This could lead to an enhanced characterization of water surface dynamics.

Another possibility to explore is the development of a spectral method aiming at detecting new SB with Sentinel-2 time series.

Bibliography

- Al-doski Jwan, Mansor Shattri B. and Shafri Helmi Zulhaidi Mohd. (2013). Change Detection Process and Techniques. *Civil and Environmental Research*, Vol.3, No.10.
- Alesheikh, A.A. ; Ghorbanali, A. ; Nouri, N. (2007). Coastline change detection using remote sensing. *Int. J. Environ. Sci. Technol*, 4, 61–66.
- Althuwaynee O.F., Pradhan B., Lee S. (2016). A novel integrated model for assessing landslide susceptibility mapping using CHAID and AHP pair-wise comparison. *Int. J. Remote Sens*, 37 (5), pp. 1190-1209 .
- Andres L., Salas W., and Skole D. (1994). “Fourier analysis of multi-temporal AVHRR data applied to a land cover classification,” *International Journal of Remote Sensing*, vol. 15, no. 5, pp. 1115–1121.
- Ballard D, Brown C. (1982). *Computer Vision*. Prentice-Hall : Englewood Cliffs, NJ.
- Banko Gebhard. (1998). A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data and of Methods Including Remote Sensing Data in Forest Inventory. INTERIM REPORT IR-98-081 / November 1998.
- Batista, G.E.A.P.A., Keogh, E.J., Tataw, O.M., de Souza, V.M.A. (2014). CID : an efficient complexity-invariant distance for time series. *Data Min. Knowl. Discov.* 28 (3), 634–669.
- Bergue J-M and Ruperd Y. (coordonnateurs) (1994). *Guide technique des bassins de retenue d’eaux pluviales*. Technique et Documentation - Lavoisier, 1994, Paris (France).
- Bertels Luc, Smets Bruno and Wolfs Davy. (2016). Dynamic Water Surface Detection Algorithm Applied on PROBA-V Multispectral Data, *Remote Sens.* 8(12), 1010 ; doi :10.3390/rs8121010.
- Bisquert, M. ; Bégué, A. ; Deshayes, M. (2015). Object-based delineation of homogeneous landscape units at regional scale based on MODIS time series. *Int. J. Appl. Earth Obs. Geoinf.* 37, 72–82.
- Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS J. Photogramm. Rem. Sens.* 65 (1), 2–16.
- Blaschke, T. (2005). Towards a framework for change detection based on image objects. *Göttinger Geogr. Abh.* 113, 1–9.
- Breiman Leo. (2001). Random forests. *Machine learning*. Vol. 45, no 1, p. 5-32.
- Byoung Chul Ko, Hyeong Hun Kim and Jae Yeal Nam. (2015). Classification of Potential Water Bodies Using Landsat 8 OLI and a Combination of Two Boosted Random Forest Classifiers. *Sensors*. 15. 13763-13777. Doi :10.3390/s150613763.
- Caetano Mario for ESA. (2007). Advanced Training course on Land Remote Sensing : Image classification. Available online at <http://earth.esa.int/landtraining07/D2L4-Caetano.pdf> (consulted 19 January 19).
- Celik T. (2009). “Unsupervised multiscale change detection in multitemporal synthetic aperture radar images,” in *17th European Signal Processing Conference (EUSIPCO)*, vol. 15, pp. 47–51.
- Celik T. and Ma K.-K. (2011). “Multitemporal image change detection using undecimated discrete wavelet transform and active contours,” *IEEE Transactions on Geoscience and Remote Sensing*, vol.

49, no. 2, pp. 706–716.

Cohen, W.N., Fiorella, M., Gray, J., Helmer, E. and Anderson K. (1998). An efficient and accurate method for mapping forest clear cuts in the Pacific Northwest using Landsat imagery. *Photogramm. Eng. Remote Sens.* 64, pp. 293- 300.

Congalton, R. G. (1991), A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* 37 :35-46.

Coppin P., Jonckheere I., Nackaerts K., Muys B., and Lambin E. (2004). Digital change detection methods in ecosystem monitoring : A review. *Int. J. Remote Sens.* Vol. 25. No. 9. Pages. 1565–1596.

Coppin P. and Bauer M. (1996). Digital change detection in forest ecosystems with remote sensing imagery. *Remote Sensing Reviews.* Vol 13 No 3-4. Pages 207-234.

Crist, E. P. and Ciconel R. C.. (1984). A physically based transformation of Thematic Mapper data—The TM Tasseled Cap. *IEEE Trans. Geosci. Remote Sens.* GE-22. 256–263.

Crist, E. P. (1985). A TM Tasseled Cap equivalent transformation for reflectance factor data. *Remote Sensing of Environment*, 17(3), 301–306. [https://doi.org/10.1016/0034-4257\(85\)90102-6](https://doi.org/10.1016/0034-4257(85)90102-6)

D’Andrimont Raphaël, Marlier Catherine, Defourny Pierre. (2017). Hyperspatial and Multi-Source Water Body Mapping : A Framework to Handle Heterogeneities from Observations and Targets over Large Areas, *Remote Sensing*, *Remote Sens.* 9. 211.

Davranche Aurélie, Lefebvre Gaëtan, Poulin B. (2010). Wetland monitoring using classification trees and SPOT-5 seasonal time series. *Remote sensing of Environment*, 114(3) :552-562. DOI : 10.1016/j.rse.2009.10.0

Débonnaire Nicolas, Stumpf André, and Puissant Anne. (2016). Spatio-Temporal Clustering and Active Learning for Change Classification in Satellite Image Time Series. *IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING.* Vol 9. Issue 8. Pages 3642 – 3650. DOI : 10.1109/JSTARS.2016.2525940.

Devadas, R. ; Denham, R.J. ; Pringle, M. (2012). Support vector machine classification of object-based data for crop mapping, using multi-temporal landsat imagery. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* XXXIX-B7, 185–190.

Domenikiotis, C., Loukas, A., Dalezios, N. R. (2003). The use of NOAA/AVHRR satellite data for monitoring and assessment of forestfires and flood. *Natural Hazards and Earth System Sciences*, 3, 115–128.

Du Z., Linghu B., Ling F., Li W., Tian W., Wang H., Gui Y., Suz B. and Zhang X. (2012). Estimating surface water area changes using time-series Landsat data in the qingjiang river basin, China. *J. Appl. Remote Sens.* 6, doi :10.1117/1.JRS.6.063609.

Dronova, I. ; Gong, P. ; Wang, L. ; Zhong, L. (2015). Mapping dynamic cover types in a large seasonally flooded wetland using extended principal component analysis and object-based classification. *Remote Sens. Environ.* 158, 193–206.

Dronova, I. ; Gong, P. ; Wang, L. (2011) Object-based analysis and change detection of major wetland cover types and their classification uncertainty during the low water period at Poyang Lake, China. *Remote Sens. Environ.* 115, 3220–3236.

Dwivedi, R. S., Kandrika, S. (2005). Delineation and monitoring of aquaculture areas using multi-

temporal space-borne multispectral data. *Current Science*, 89 (8), 1414–1421.

Feng Quanlong, Liu Jiantao and Gong Jianhua. (2015). Urban flood mapping based on unmanned aerial vehicle remote sensing and random forest classifier—A case of Yuyao, China. *Water*. Vol. 7, no 4, p. 1437-1455.

Feyisa, G. L., Meilby, H., Fensholt, R., Proud, S. R. (2014). Automated water extraction index : A new technique for surface water mapping using Landsat imagery. *Remote Sensing of Environment*, 140(0), 23–35. <https://doi.org/10.1016/j.rse.2013.08.029>.

Fisher Adrian and Danaher Tim. (2013), A Water Index for SPOT5 HRG Satellite Imagery, New South Wales, Australia, Determined by Linear Discriminant Analysis, *Remote sensing*. 2013, 5, 5907-5925 ; doi :10.3390/rs5115907.

Fisher, A., Flood, N., Danaher, T. (2016). Comparing Landsat water index methods for automated water classification in eastern Australia. *Remote Sensing of Environment*, 175, 167–182. <https://doi.org/10.1016/j.rse.2016.05.013>.

Foody, G. M. (1992). On the compensation for chance agreement in image classification accuracy assessment. *Photogrammetric Engineering and Remote Sensing*, 58, 1459–1460.

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*. 80, 185–201.

Frazier Paul Shane and Page Kenneth John. (2000). Water body detection and Delineation with Landsat TM data. *Photogrammetric Engineering Remote Sensing*. Volume 66, Number 12, Pages 1462-1467.

Friedl M A and Brodley C E. (1997). Decision tree classification of land cover from remotely sensed data ; *Remote Sens. Environ.* 61 399–409.

Frohn, R. ; Autrey, B. ; Lane, C. ; Reif, M. (2011). Segmentation and object-oriented classification of wetlands in a Karst Florida landscape using multi-season Landsat-7 ETM+ imagery. *Int. J. Remote Sens.* 32, 1471–1489.

Ghimire, B., Rogan, J., Rodriguez Galiano, V., Panday, P., Neeti, N. (2012). An evaluation of bagging, boosting, and Random Forests for land cover classification in Cape Cod, Massachusetts, USA. *GIScience Remote Sens.* 49 (5), 623–643.

Guo, M., Li, J., Sheng, C., Xu, J., and Wu, L. (2017). A Review of Wetland Remote Sensing. *Sensors*. 17. 777. URL : <https://doi.org/10.3390/s17040777>.

Gusmawati Niken, Soulard Benoît, Selmaoui-Folcher Nazha, Proisy Christophe, Mustafa Akhmad, Le Gendre Romain, Laugier Thierry, Lemonnier Hugues. (2017). Surveying shrimp aquaculture pond activity using multitemporal VHSR satellite images - case study from the Perancak estuary, Bali, Indonesia. *Marine Pollution Bulletin*. N 08520. Page 12.

Guttler Fabio, Ienco Dino, Nin Jordi, Teisseire Maguelonne and Poncelet Pascal. (2017). A graph-based approach to detect spatiotemporal dynamics in satellite image time series. *ISPRS Journal of Photogrammetry and Remote Sensing*. 130. 92–107. <http://dx.doi.org/10.1016/j.isprsjprs.2017.05.013>.

Guttler Fabio., Alleaume S., Corbane C., Ienco D., Nin Jordi, Poncelet Pascal, Teisseire Maguelonne. (2014). Exploring high repetitivity remote sensing time series for mapping and monitoring natural habitats - A new approach combining OBIA and k-partite graphs. In : *IGARSS*.

- Haas J., Ban Y. (2017). Sentinel-1A SAR and sentinel-2A MSI data fusion for urban ecosystem service mapping. *Remote Sens. Appl. Soc. Environ.* 8 :41–53. doi : 10.1016/j.rsase.2017.07.006.
- Hassan, N., Hamid, J.R.A., Adnan, N.A. and Jaafar, M. (2014) Delineation of Wetland Areas from High Resolution World View-2 Data by Object-Based Method. *Proceedings of the IOP Conference Series : Earth and Environmental Science*, 18, Article ID : 012017.
- Huang Chang, Chen Yun, Zhang Shiqiang, Wu Jianping. (2018). Detecting, Extracting, and Monitoring Surface Water From Space Using Optical Sensors : A Review. *Reviews of Geophysics*, Volume 56, Issue 2.
- Huang, C., Thomas, N., Goward, S. N., Masek, J. G., Zhu, Z., Townshend, J. R., and Vogelmann, J. E. (2010). Automated masking of cloud and cloud shadow for forest change analysis using Landsat images. *International Journal of Remote Sensing*, 31(20) :5449–5464.
- Irons James R., Dwey John L. and Barsi Julia A. (2012). The next Landsat satellite : The Landsat Data Continuity Mission. *Remote Sensing of Environment*, Volume 122, Pages 11-21.
- Ivanovsky, A., Belles, A., Criquet, J., Dumoulin, D., Noble, P., Alary, C., Billon, G. (2018). Assessment of the treatment efficiency of an urban stormwater pond and its impact on the natural downstream watercourse. *Journal of Environmental Management*, 226, 120–130. doi :10.1016/j.jenvman.2018.08.015.
- Jain, S. K., Saraf, A. K., Goswami, A., Ahmad, T. (2006). Flood inundation mapping using NOAA AVHRR data. *Water Resources Management*, 20, 949–959.
- Jain, S. K., Singh, R. D., Jain, M. K., Lohani, A. K. (2005). Delineation of flood-prone areas using remote sensing technique. *Water Resources Management*, 19, 337–347.
- Javaux Mathieu ; Delforge Damien Dubois, Damien. (2014). Contrat de gestion des risques d'inondation liés aux crues et ruissellements en Brabant Wallon. Published Online at : <http://hdl.handle.net/2078.1/156139>.
- Jawak Shridhar D., Devliyal Prapti and Luis Alvarinho J. (2015). A Comprehensive Review on Pixel Oriented and Object Oriented Methods for Information Extraction from Remotely Sensed Satellite Images with a Special Emphasis on Cryospheric Applications. *Advances in Remote Sensing*. 4. 177-195. Published Online in SciRes. <http://www.scirp.org/journal/ars> <http://dx.doi.org/10.4236/ars.2015.43015>.
- Jayanthi, M. (2011). Monitoring brackishwater aquaculture development using multi-spectral satellite data and GIS : A case study near Pichavaram mangroves south-east coast of India. *Indian Journal of Fisheries*, 58 (1), 85.
- Ji, L., Zhang, L., Wylie, B. (2009). Analysis of dynamic thresholds for the normalized difference water index. *Photogrammetric Engineering and Remote Sensing*, 75(11), 1307–1317.
- Ji, L. Y., Geng, X. R., Sun, K., Zhao, Y. C., Gong, P. (2015). Target detection method for water mapping using Landsat 8 OLI/TIRS imagery. *Water*, 7(2), 794–817. <https://doi.org/10.3390/w7020794>.
- Jiang, H., Feng, M., Zhu, Y., Lu, N., Huang, J., Xiao, T. (2014). An Automated Method for Extracting Rivers and Lakes from Landsat Imagery. *remote sensing*, 6, 5067-5089. <https://doi.org/10.3390/rs6065067>.
- Julea A. et al. (2011). “Unsupervised spatiotemporal mining of satellite image time series using grouped frequent sequential patterns,” *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 4, pp. 1417–1430.
- Kallio, K., J. Attila, P. Härmä, S. Koponen, J. Pulliainen, U. -M. Hyytiäinen, and T. Pyhälähti. (2008). Landsat ETM+ Images in the estimation of Seasonal lake water quality in Boreal rivers basins.

Environ. Manage. 42 :511-522 [doi :10.1007/ s00267-008-9146-y].

Kaplan, G., Avdan, U. (2018). SENTINEL-1 AND SENTINEL-2 DATA FUSION FOR WETLANDS MAPPING : BALIKDAMI, TURKEY. *International Archives of the Photogrammetry, Remote Sensing Spatial Information Sciences*, 42(3).

Kawakubo, F. S., Morato, R. G., Nader, R. S., Luchiari, A. (2011). Mapping changes in coastline geomorphic features using Landsat TM and ETM+imagery : examples in south- eastern Brazil. *International Journal of Remote Sensing*, 32(9), 2547–2562. doi :10.1080/01431161003698419.

Klein, I., Dietz, A., Gessner, U., Dech, S., and Kuenzer, C. (2015). Results of the Global WaterPack : a novel product to assess inland water body dynamics on a daily basis, *Remote Sens. Lett.*, 6, 78–87, <https://doi.org/10.1080/2150704X.2014.1002945>.

Kux, H.J. and Souza, U.D. (2012). Object-Based Image Analysis of WORLDVIEW-2 Satellite Data for the Classifica- tion of Mangrove Areas in the City of SÃO LUÍS, MARANHÃO State, Brazil. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1, 95-100. <http://dx.doi.org/10.5194/isprsannals-I-4-95-2012>.

Li, J. L., Sheng, Y. W. (2012). An automated scheme for glacial lake dynamics mapping using Landsat imagery and digital elevation models : A case study in the Himalayas. *International Journal of Remote Sensing*, 33(16), 5194–5213. <https://doi.org/10.1080/01431161.2012.657370>.

Lillesand, Thomas M., Kiefer, Ralph W., Chipman, Jonathan W. (2015). *Remote sensing and image interpretation*. 7th ed, Wiley Sons, New York.

Lin, Li Di, Liping Genong Yu, Eugene Kang, Lingjun Shrestha, Ranjay Rahman, Md Tang, Junmei Deng, Meixia Sun, Ziheng Zhang, Chen Hu, Lei. (2016). A review of remote sensing in flood assessment. 1-4. 10.1109/Agro-Geoinformatics.2016.7577655.

Liu H, Wang L. (2008). Mapping detention basins and deriving their spatial attributes from airborne LiDAR data for hydrological applications. *Hydrological Processes* 22 : 2358–2369.

Longbotham, N. ; Pacifici, F. ; Glenn, T. ; Zare, A. ; Volpi, M. ; Tuia, D. ; Christophe, E. ; Michel, J. ; Inglada, J. ; Chanussot, J. ; et al. (2012). Multi-modal change detection, application to the detection of flooded areas : Outcome of the 2009–2010 data fusion contest. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 5, 331–342.

Lopez-Caloca,A.;Tapia-Silva,F.O.;Escalante-Ramirez,B. (2008). Lake Chapala change detection using time series. *Remote Sens. Agric. Ecosyst. Hydrol.* 7104, 1–11.

Lu D., Mausel P., Brondízio E., Moran E. (2004). Change detection techniques. *International Journal of Remote Sensing* 25, 2365–2401.

Ma, L., Li, M., Ma, X., Cheng, L., Du, P., Liu, Y. (2017). A review of supervised object-based land-cover image classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 130, 277–293. doi :10.1016/j.isprsjprs.2017.06.001.

Ma Shengfang, Zhouc Yuting, Gowdad Prasanna H., Donge Jinwei, Zhangf Geli, Kakanic Vijaya G., Wagled Pradeep, Chena Liangfu, Flynnng K. Colton and Jiangh Weiguo. (2019). Application of the water-related spectral reflectance indices : A review. *Ecological Indicators*.Volume 98, Pages 68-79.

Ma Z. , Redmond R.L. (1995). Tau coefficients for accuracy assessment of classification of remote

sensing data. *Photogramm. Eng. Remote Sens.* 61 , pp. 435-439.

Mahé, G., Orange, D., Mariko, A., and Bricquet, J. P. (2011). Estimation of the flooded area of the Inner Delta of the River Niger in Mali by hydrological balance and satellite data, *IAHS-AISH P.*, 344, 138–143.

Mallinis, G., Gitas, I. Z., Giannakopoulos, V., Maris, F., Tsakiri-Strati, M. (2011). An object-based approach for flood area delineation in a transboundary area using ENVISAT ASAR and LANDSAT TM data. *International Journal of Digital Earth*, 1–13. doi :10.1080/17538947.2011.641601

Malinowski, R., Groom, G., Schwanghart, W., Heckrath, G. (2015). Detection and delineation of localized flooding from WorldView-2 multispectral data. *Remote Sensing*, 7(11), 14853.

Manavalan, P., P. Sathyanath, and G. L Rajegowda. (1993). Digital image analysis techniques to estimate waterspread for capacity evaluations of reservoirs. *Photogramm. Eng. Remote Sens.* 59(9) :1389-1395.

Marjanović M., Kovačević M., Bajat B., Voženílek V. (2011) Landslide susceptibility assessment using SVM machine learning algorithm. *Eng. Geol.*, 123 (3), pp. 225-234

McFeeters, S. K. (1996). The use of the normalized difference water index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17(7), 1425–1432.

Mougel P. and Selmaoui-Folcher N. (Jul. 2013). “A data mining approach to discover collections of homogeneous regions in satellite image time series,” in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, pp. 4360–4363.

Mountrakis, G., Im, J. and Ogole, C. (2011). Support Vector Machines in Remote Sensing : A Review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66, 247-259. <http://dx.doi.org/10.1016/j.isprsjprs.2010.1>

Nath R.K, Deb S.K. (2010). Water-Body Area Extraction from High Resolution Satellite Images—An Introduction, Review, and Comparison. *International Journal of Image Processing*, 3, 353-372.

Ogilvie Andrew, Belaud Gilles, Massuel Sylvain, Mulligan Mark, Le Goulven Patrick, and Calvez Roger. (2018). Surface water monitoring in small water bodies : potential and limits of multi-sensor Landsat time series. *Hydrology and Earth System Sciences* 22, 4349–4380.

Overton, I. (1997). Satellite image analysis of river Murray floodplain inundation. NRMS Project R6045, Murra-Darling Basin Commission.

Papa, F., Prigent, C., Aires, F., Jimenez, C., Rossow, W. B., and Matthews, E. : Interannual variability of surface water extent at the global scale, 1993–2004, *J. Geophys. Res.*, 115, D12111, <https://doi.org/10.1029/2009JD012674>, 2010.

Paul, A., Tripathi, D., Dutta, D. (2017). Application and comparison of advanced supervised classifiers in extraction of water bodies from remote sensing images. *Sustainable Water Resources Management*. doi :10.1007/s40899-017-0184-6

Pekel Jean-François, Cottam Andrew, Gorelick Noel, S. Belward Alan. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*. Volume 540, Pages 418–422.

Petitjean F., Kurtz C., Passat N., Gançarski P. (2012b). Spatio-temporal reasoning for the classification of satellite image time series. *Pattern Recogn. Lett.* 33 (13), 1805–1815.

- Petitjean F., Inglada J., Gançarski P. (2012a). Satellite image time series analysis under time warping. *IEEE Trans. Geos. and Rem. Sensing* 50 (8), 3081–3095.
- Petitjean F., Gançarski P., Masseglia F., and Forestier G. (2010) “Analysing satellite image time series by means of pattern mining,” in *Intelligent Data Engineering and Automated Learning—IDEAL*, C. Fyfe, P. Tino, D. Charles, C. Garcia-Osorio, and H. Yin, Eds. New York, NY, USA : Springer, vol. 6283, pp. 45–52.
- Pontius, R. G. (2000). Quantification error versus location error in comparison of categorical maps. *Photogrammetric Engineering and Remote Sensing*, 66, 1011–1016.
- Pontius Jr. R. G., Mollines M. (2011). Death to Kappa : birth of quantity disagreement and allocation disagreement for accuracy assessment. *Int. Journal of Remote Sensing*, Vol. 32, No. 15, 4407–4429.
- Poulain D., Touze-Foltz N., Peyras L., Duquennoi C. (2011). Containment ponds, reservoirs and canals. In *Handbook of Geosynthetic Engineering*. ICE, 32 pages. ISBN : 978-0-7277-4175-2.
- Prigent, C., Papa, F., Aires, F., Rossow, W. B., and Matthews, E. (2007). Global inundation dynamics inferred from multiple satellite observations, 1993–2000, *J. Geophys. Res.*, 112, D12107, <https://doi.org/10.1029/2006JD007847>.
- Puech, C. (1994). Plans d’eau sahéliens et imagerie SPOT : inventaire et évaluation des capacités d’exploitation. Colloque international “Eau, environnement et développement” Colloque de l’Institut des Relations Interuniversitaires avec la Mauritanie (IRIM)-Université de Nice/Sophia-Antipolis et de l’Institut Mauritanien de la Recherche Scientifique (IMRS), Nouakchott, 20–22 mars 1994 (pp. 6883).
- Pulvirenti, L., Marzano, F. S., Pierdicca, N., Mori, S., Chini, M. (2013). Discrimination of water surfaces, heavy rainfall, and wet snow using COSMO-SkyMed observations of severe weather events. *IEEE transactions on geoscience and remote sensing*, 52(2), 858–869.
- Quintano, C., Fernández-Manso, A., Fernández-Manso, O., Shimabukuro, Y. E. (2006). Mapping burned areas in Mediterranean countries using spectral mixture analysis from a uni-temporal perspective. *International Journal of Remote Sensing*, 27(4), 645–662. doi :10. 1080/01431160500212195.
- Quintano, C., Shimabukuro, Y. E., Fernández, A., Delgado, J. A. (2005). A spectral unmixing approach for mapping burned areas in Mediterranean countries. *International Journal of Remote Sensing*, 26(7), 1493–1498. doi :10. 1080/01431160412331330220.
- Radoux, Julien ; Chomé, Guillaume ; Jacques, Damien C. ; Waldner, François ; Bellemans, Nicolas ; Matton, Nicolas ; Lamarche, Céline ; D’Andrimont, Raphaël ; Defourny, Pierre. (2016). "Sentinel-2’s Potential for Sub-Pixel Landscape Feature Detection." *Remote Sens.* 8, no. 6 : 488.
- Radoux, J. (2010). Updating land cover maps by GIS-driven analysis of very high spatial resolution satellite images. Ph.D. thesis, UCL.
- Rejichi S. and Chaabane F. (Jul. 2012) “Pixel and region based temporal classification fusion for HR satellite image time series,” in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, pp. 435–438.
- Rizzi, R., Rudorff, B. F. T., Shimabukuro, Y. E., Doraiswamy, P. C. (2006). Assessment of MODIS LAI retrievals over soybean crop in Southern Brazil. *International Journal of Remote Sensing*, 27(19), 4091–4100. doi :10.1080/01431160600851850.

- Rodríguez Galiano, V.F., Chica Olmo, M., Abarca Hernández, F., Atkinson, P.M., Jeganathan, C. (2012). Random forest classification of mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Remote Sens. Environ.* 121, 93–107.
- Rokni, K., Ahmad, A., Selamat, A., Hazini, S. (2014). Water feature extraction and change detection using multitemporal Landsat imagery. *Remote Sensing*, 6(5), 4173–4189.
- Romani L. et al. (Jan. 2013). “A new time series mining approach applied to mul- titemporal remote sensing imagery,” *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 140–150.
- Rosenfield, G. H., Fitzpatrick-Lins, K. (1986). A coefficient of agree- ment as a measure of thematic classification accuracy. *Photogrammetric Engineering and Remote Sensing*, 52, 223–227.
- Sawaya Kali E. , Olmanson Leif G., Heinert Nathan J., Brezonik Patrick L., Bauer Marvin E. (2003). Extending satellite remote sensing to local scales : land and water resource monitoring using high-resolution imagery *Remote Sensing of Environment* 88 (2003) 144–156
- Sarp, G., Ozcelik, M. (2017). Water body extraction and change detection using time series : A case study of Lake Burdur, Turkey. *Journal of Taibah University for Science*, 11(3), 381–391. doi :10.1016/j.jtusci.2016
- Sharma Richa, Ghosh Aniruddha and Joshi P K. (2013). Decision tree approach for classification of remotely sensed satellite data using open source support. *Journal Earth System Science*. Volume 122. Issue 5. Pages 1237–1247. doi.org/10.1007/s12040-013-0339-2
- Shimabukuro, Y. E., Duarte, V., Arai, E., et al. (2009). Fraction images derived from Terra Modis data for mapping burnt areas in Brazilian Amazonia. *International Journal of Remote Sensing*, 30(6), 1537–1546. doi :10.1080/01431160802509058.
- Singh Y., Kaur A., Malhotra R. (2009). Comparative analysis of regression and machine learning methods for predicting fault proneness models. *Int. J. Comput. Appl. Technol.* 35 (2–4), pp. 183-193.
- Singh A. (1989). Review Article Digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing*. Vol 43. Issue 6. Pages 989-1003, 1989. DOI 10.1016/0034-4257(79)90013-0
- Singh, A. (1986). Change detection in the tropical forest environment of northeastern India using Landsat, Remote sensing and tropical land management. J. Wiley, New York, pp. 237–254.
- Sivanpillai Ramesh and Miller Scott N. (2010). Improvements in mapping water bodies using ASTER data. *Ecological Informatics* 5(1) :73-78. DOI : 10.1016/j.ecoinf.2009.09.013
- Smith, L. C. (1997). Satellite remote sensing of river inundation area, stage and discharge : a re- view. *Hydrolog. Proc.* 11 :1427-1439 [doi :10.1002/(SICI)1099-1085(199708)11 :10 <1427 : :AID-HYP473>3.0.CO ;2-S].
- Smits, P. C., Dellepiane, S. G., Schowengerdt, R. A. (1999). Quality assessment of image classifica- tion algorithms for land-cover mapping : a review and proposal for a cost-based approach. *Internatio- nal Journal of Remote Sensing*, 20, 1461–1486.
- Song, C., C. E. Woodcock, K. C. Seto, M. P. Lenney, and S. A. Macomber. (2002). Classification and change detection using Landsat TM data : When and how to correct atmospheric effects ? *Remote Sens. Environ.* 75 :230-244 [doi :10.1016/ S0034-4257(00)00169-3].
- Soti V., Puech C., Lo Seen D., Bertran A., Vignolles C., Mondet B., Dessay N. and Tran A. (2010),

- The potential for remote sensing and hydrologic modelling to assess the spatio-temporal dynamics of ponds in the Ferlo Region (Senegal). *Hydrology and Earth System Sciences*. 14 : 1449–1464.
- Sridhar, P. N., Surendran, A., Ramana, I. V. (2008). Auto extraction technique based digital classification of salt pans and aquaculture plots using satellite data. *International Journal of Remote Sensing*, 29 (2), 313–323. doi :10.1080/014311607012503
- Starzec Peter, Lind BO B., Lanngren Anders, Lindgren Åsa, Svenson Torbjörn. (May 2005). Technical and Environmental Functioning of Detention Ponds for the Treatment of Highway and Road Runoff. *Water, Air, and Soil Pollution*. 163 : 1–4. Pages 153–16.
- Stehman Stephen. (1997). Selecting and interpreting measures of thematic classification accuracy. *Remote Sens. Environ*. 62 :77-89.
- Su L., Huang, Y. (2009). Support vector machine (SVM) classification : comparison of linkage techniques using a clustering-based method for training data selection. *GIScienceRemoteSensing*46(4),411–423.
- Swenson, S. and Wahr, J. (2009). Monitoring the water balance of Lake Victoria, East Africa, from space, *J. Hydrol.*, 370, 163–176, <https://doi.org/10.1016/j.jhydrol.2009.03.008>.
- Tao Zhang, Xiaomei Yang, Shanshan Hu, and Fenzhen Su. (2013). Extraction of Coastline in Aquaculture Coast from Multispectral Remote Sensing Images : Object-Based Region Growing Integrating Edge Detection, *Remote Sens*. 5 : 4470-4487.
- Tehrany Mahyat Shafapour, Pradhan Biswajeet and Jebur Mustafa Neamah. (2013). Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. *Journal of Hydrology*. 504. DOI : 10.1016/j.jhydrol.2013.09.034.
- Tien Bui D., Tuan T.A., Klempe H., Pradhan B., Revhaug I., G. (2016). Spatial prediction models for shallow landslide hazards : a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides*, 13 (2), pp. 361-378
- Tien Bui D., Pradhan B., Lofman O., Revhaug I. Landslide susceptibility assessment in Vietnam using support vector machines, decision tree, and Naive Bayes Models *Math. Probl. Eng.*, 2012 (2012)
- Travaglia, C., Profeti, G., Aguilar-Manjarrez, J., Lopez, N. A. (2004). Mapping coastal aquaculture and fisheries structures by satellite imaging radar : Case study of the Lingayen Gulf, the Philippines (FAO Fisheries Technical Paper).Rome :Food
- Turk, G. (1979). GT index : a measure of the success of prediction. *Remote Sensing of Environment*, 8, 75–86.
- Tyralis H., Papacharalampous G. and Et Langousis A. (2019). A brief review of random forests for water scientists and practitioners and their recent history in water resources. *Water*. Vol. 11, no 5, p. 910.
- Verpoorter, C., T. Kutser, and L. Tranvik. (2012). Automated mapping of water bodies using Landsat multispectral data. *Limnol. Oceanogr. Methods*. 10. 1037–1050. doi :10.4319/lom.2012.10.1037.
- Virdis Salvatore Gonario Pasquale. (2013). An object-based image analysis approach for aquaculture ponds precise mapping and monitoring : a case study of Tam Giang-Cau Hai Lagoon, Vietnam. *Environ Monit Assess*.186 :117–133. DOI 10.1007/s10661-013-3360-7.

- Wang L, Liu H. (2006). An efficient method for identifying and filling surface depressions in digital elevation models for hydrological analysis and modeling. *International Journal of Geographical Information Science* 20 : 193 – 213.
- Wang L., Yu J. (2012). Modelling detention basins measured from high-resolution light detection and ranging data. *Hydrological Processes*. 26 :19. Pages 2973-2984.
- Wang, X. ; Xie, H. (2018). A Review on Applications of Remote Sensing and Geographic Information Systems (GIS) in Water Resources and Flood Risk Management. *Water*, 10, 608.
- Warrens, Matthijs. (2015). Five Ways to Look at Cohen's Kappa. *Journal of Psychology Psychotherapy*. 05. 10.4172/2161-0487.1000197.
- White, K., and H. M. El Asmar. (1999). Monitoring changing position of coastlines using thematic mapper imagery, an example from the Nile Delta. *Geomorphology* 29 :93-105 [doi :10.1016/S0169-555X(99)00008-2].
- Wolski, P. and Murray-Hudson, M. (2008). An investigation of permanent and transient changes in flood distribution and outflows in the Okavango Delta, Botswana, *Phys. Chem. Earth Pt. A/B/C*, 33, 157–164, <https://doi.org/10.1016/j.pce.2007.04.008>.
- Wu, Q.Z., Cheng, H.Y., Jeng, B.S., (2005). Motion detection via change-point detection for cumulative histograms of ratio images. *Pattern Recognition Lett.* 26 (5), 555–563.
- Wu X., Ren F., Niu R. (2014). Landslide susceptibility assessment using object mapping units, decision tree, and support vector machine models in the Three Gorges of China. *Environ. Earth Sci.*, 71 (11), pp. 4725-4738.
- Xu H. Q. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, 27(14), 3025–3033.
- Xu Y.B. ; Lai, X.J. ; Zhou, C.G. (2010). Water surface change detection and analysis of bottomland submersion-emersion of wetlands in Poyang Lake reserve using ENVISAT ASAR data. *China Environ. Sci.* 30, 57–63.
- Xue C. J., Dong Q., and Ma W. X. (2014). “Object-oriented spatial-temporal association rules mining on ocean remote sensing imagery,” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 17, no. 1, p. 012109.
- Yao, F., Wang, C., Dong, D., Luo, J., Shen, Z., Yang, K. (2015). High-resolution mapping of urban surface water using ZY-3 multi-spectral imagery. *Remote Sensing*, 7(9), 12336.
- Yu Q., Gong P., Chinton N., Biging G., Kelly M. and Schirokauer D. (2006). Object- based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. *Photogrammetric Engineering Remote Sensing* 72 (7), 799–811.
- Zhang, H., Li, J., Xiang, N., Shen, Q., Zhang, F., Liang, W. (2014). An automatic method of monitoring water bodies based on GF-1 data. Paper presented at SPIE Ocean Remote Sensing and Monitoring from Space, Beijing, China.
- Zhang, T., Yang, X., Hu, S., Su, F. (2013). Extraction of Coastline in Aquaculture Coast from Multispectral Remote Sensing Images : Object-Based Region Growing Integrating Edge Detection. *Remote Sensing*, 5(9), 4470–4487. doi :10.3390/rs5094470
- Zhang, T., Li, Q., Yang, X., Zhou, C., Su, F. (2010). Automatic mapping aquaculture in coastal zone

from TM imagery with OBIA approach. In *Geoinformatics , 2010 18th International Conference on Geoinformatics*, 18–20 June 2010 (pp. 1–4)

Zhou Yan , Dong Jinwei, Xiao Xiangming, Xiao Tong, Yang Zhiqi, Zhao Guosong, Zou Zhenhua and Qin Yuanwei. (2017). Open Surface Water Mapping Algorithms : A Comparison of Water-Related Spectral Indices and Sensors. *Water*. 9(4). 256. doi :10.3390/w9040256 Zhou, H. ; Hong, J. ; Huang, Q. Landscape and water quality change detection in urban wetland : A post-classification comparison method with IKONOS data. *Procedia Environ. Sci.* 2011, 10, 1726–1731.

Zhu, Q., Li, J., Zhang, F., Shen, Q. (2017). Distinguishing Cyanobacterial Bloom From Floating Leaf Vegetation in Lake Taihu Based on Medium-Resolution Imaging Spectrometer (MERIS) Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(1), 34-44.

Zhu, Z., Wang, S., and Woodcock, C. E. (2015). Improvement and expansion of the Fmask algorithm : Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sensing of Environment*, 159 :269–277.

Zhu, Q., Batista, G.E.A.P.A., Rakthanmanon, T., Keogh, E.J. (2012). A novel approximation to dynamic time warping allows anytime clustering of massive time series datasets. In : *SDM*, pp. 999–1010.

Zhu, G. and Blumberg, D. G. (2002). Classification using ASTER data and SVM algorithms ; The case study of Beer Sheva, Israel. *Remote Sensing of Environment*, 80, 233-240.

Webography

Walloon Brabant Province. (2019a). Website. "Plateforme provinciale de gestion des risques d'inondation". Available at : <http://www.brabantwallon.be/inondations/> (Assessed 2 June 2019).

Walloon Brabant Province. (2019b). Website. "Historique des inondations". Available at : <http://www.brabantwallon.be/inondations/> (Assessed 2 June 2019).

Bruxelles environment. (2010a) Website. "Fiche informative outil de gestion de l'eau de pluie OGE02 – Bassin sec". Available at <https://www.guidibatimentdurable.brussels/servlet/Repository/iframe-eco-construction-oge02-bassin-sec.pdf?ID=34298&fbclid=IwAR3291RrujZDrUYUUPOcuiXNvapru2elcrpSMfvkXf9ZyNsXd6YDQjT9rTQ> (Assessed 22 January 2019).

Bruxelles environment. (2010b) Website. "Fiche informative outil de gestion de l'eau de pluie OGE03 – Bassin en eau". Available at : https://www.guidibatimentdurable.brussels/servlet/Repository/iframe-eco-construction-oge02-bassin-en-eau.pdf?ID=34299&saveFile=true&fbclid=IwAR1E3eei197Sc9SFQUzhExXmiFfBdJApIoD-Uk7DQ-tvJ_Sh0D-4p79eSV4 (Assessed 22 January 2019).

Code de l'Eau de la Wallonie. Available at : <http://environnement.wallonie.be/legis/menucode.htm>. (Assessed on 2 June 2018).

Directive 2007/60/CE du Parlement Européen et du Conseil du 23 octobre 2007 relative à l'évaluation et à la gestion des risques d'inondation. Available at : <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX:32007L0060>. (Assessed on 2 June 2018).

European Space Agency. (2019a). Website. "Sentinel-2 mission Description". Available at : <https://sentinel.esa.int/web/sentinel/missions/sentinel-2> (Assessed 2 June 2019).

European Space Agency. (2019b). Website. "MultiSpectral Instrument (MSI) Overview". Available at : <https://earth.esa.int/web/sentinel/technical-guides/sentinel-2-msi/msi-instrument> (Assessed 2 June 2019).

European Space Agency and Sentinel-2 for Agriculture. (2018). Website. "Detailed description of the Sen2-Agri system". <http://www.esa-sen2agri.org/operational-system/system-description> (Assessed 2 June 2019).

Federal Public Service Interior. (2019). Website. "Calamités naturelles – Fonds des Calamités fédéral". Available at : <https://www.ibz.be/fr/calamites-naturelles-fonds-des-calamites>. (Assessed 2 June 2018).

Federal Public Service Justice. (2019). Website. "Moniteur Belge". Available at : http://www.ejustice.just.fgov.be/doc/rech_f.htm. (Assessed 2 June 2018).

Geoportail of Wallonia. (2018). Website. "Orthophotos : La Wallonie vue du ciel". Available at : <http://geoportail.wallonie.be/georeferentiel/orthophotos> (Assessed 2 June 2019).

Laramie County Conservation District (LCCD). (2016). Website. "Retention vs. Detention ponds ».

Available at :<https://www.lccdnet.org/wp-content/uploads/Ponds.pdf> (Assessed 2 June 2019).

Public Service of Wallonia. (2019). "Calamités naturelles - Fonds Régional des Calamités : Actualités". Available at :<http://pouvoirslocaux.wallonie.be/jahia/Jahia/site/dgpl/accueil/calamites>. (Assessed 2 June 2018)

Public Service of Wallonia for Mobility and Infrastructure. (2019). Website. "Identifier une route régionale / route communale". Available at :<https://www.wallonie.be/demarches/138360-identifier-une-route-regionale-route-communale> (Assessed 2 June 2019).

Royal Meteorological Institute of Belgium (IRM). (2019). Website. "Le climat de la Belgique". Available at :<https://www.meteo.be/fr/climat/atlas-climatique/le-climat-de-la-belgique-contexte-du-climat-de-la-belgique>. (Assessed 2 June 2019).

Appendix A. Illustrations of SB



(a) E411, Corroy-le-Grand



(b) Pisselet, Grez-Doiceau



(c) Sablière, Mont-Saint-Guibert (Left)



(d) Sablière, Mont-Saint-Guibert (Right)

FIGURE A. 1 – Illustrations of excellent SB



(a) Athena, Louvain-la-Neuve



(b) E411 Bretelle 8A, Chaumont-Gistoux



(c) Procession aux reliques, Wavre (Left)



(d) Procession aux reliques, Wavre (Right)

FIGURE A. 2 – Illustrations of medium SB



(a) Stampia RN25, Grez-Doiceau (Remote perspective)



(b) Stampia RN25, Grez-Doiceau (Inside view)



(c) Venelle du Grand Bon Dieu du Tour, Wavre (Remote perspective)



(d) Venelle du Grand Bon Dieu du Tour, Wavre (Inside view)

FIGURE A. 3 – Illustrations of challenging SB

Appendix B. List of non-delineated SB

The boundary delineation occurred for 140 SB out of the 156 artificial open basins originally recorded in the provincial dataset. The 16 other SB recorded in the database could not be found for the following reasons.

- Three SB had already been recorded in the database while they were still at a project stage or under construction (as correctly mentioned in the database field called « operating status »)
 - Le mortier (en projet)
 - E411 – Chaumont Gistoux – Rue de Mèves (en projet)
 - Bo de l’Espinette, Tubize (en projet, trace de terres sur les ortho 2018)
- Two SB were still acknowledged to be at the project stage while they had been built and were-deconstructed ever since.
 - CET Van Gansewinkel Cour au bois, Braine-le-Château (site en remaniement)
 - Prigogyne, Mont-Saint Guibert (Renseignement pris auprès d’inBW)
- Three SB were impossible to locate even though they were recorded in the platform. For them, the provincial administration in charge of the platform update sought for clarifications from the administrations who reported those basins in order to verify their existence. At the time of the delineation, the updated information about the location of the three basins was not yet reported by the concerned authorities.
 - Avenue pasteur, wavre (Commune)
 - Rue chavée, Braine l’alleud (Commune)
 - E19 – N252 Accès Mons (SPW)
- Four SB were recorded in the platform and the province confirmed their existence but but could not be identified from orthophotoplans.
 - Cortil bailly, lasne
 - Nivelles
 - Marchal 411 - Dion
- Four other SB were impossible to delineate owing to the presence of trees.
 - Bo d’oisquercq, Tubize
 - BO rue du grand arbre, Genappe
 - BO du bois de matelle
 - Bois des gardes

This step also helped to identify and revise some outdated information of the provincial platform database. In fact, the inspection of orthophotoplans revealed that two SB were still characterized as an ongoing project even though they were constructed.

- "SB du cours d’eau » (constructed on the orthophotoplans 2016)
- « Jaurdinia » (constructed on the orthophotoplans 2006-2007)

However, the real situation can not always be depicted so easily by the orthophotoplans. As an illustration, two existing water bodies were characterized as an ongoing project. This information seemed to be outdated since the water bodies were visible, but it was not since the water bodies were in the process to be transformed in man-made SB.

- Lotissement de l'Esperance (waterbodies soon to be transformed in SB)
- « Trois fontaines » (lagooning soon to be transformed into SB)

Finally, a new basin has been located at the following address : « Chemin du Hameau, 1300 Wavre » and reported at the province administration for recording.

Appendix C. Separability Analysis

1. SFFS Algorithm and JM Distance

Dates	Band 2	Band 3	Band 4	Band 8	Band 5	Band 6	Band 7	Band 8A	Band 11	Band 12	NDWI	mNDWI	AWEIsh	AWEInsh	LSWI	mLSWI	SVM	MSI	h	s	v	JM Distance	Number of features
15/07/2018		x	x							x	x			x	x	x		x	x	x	/	1,414	10
18/09/2018		x		x		x			x	x					x	x			x	x	/	1,408	11
10/10/2018	x	x			x			x	x		x	x	x	x	x	x		x	x		/	1,414	13
3 dates		x			x		x		x	x		x	x	x	x		x	x	x	x	/	1,400	13

TABLE C. 1 – Best subsets of features obtained for the different dates by letting the algorithm freely choose among the complete set of features.

Dates	Band 2	Band 4	Band 8A	Band 11	NDWI	AWEInsh	AWEIsh	LSWI	s	JM Distance	Number of selected features	Trend relative to previous test
15/07/2018		x	x	x	x	x	x	x	x	1,410	8	-2,83%
18/09/2018	x	x	x	x	x	x	x	x		1,363	8	-3,2%
10/10/2018	x	x	x	x	x	x	x	x		1,336	8	-5,12%
3 dates		x	x	x	x	x	x	x	x	1,334	8	-4,17%

TABLE C. 2 – Best subsets of features obtained for the different dates by letting the algorithm freely choose among the following reduced set of features :Band 2, Band 4, Band 8A, Band 11, NDWI, AWEInsh, AWEIsh, LSWI and s.

Dates	Band 2	Band 4	Band 8	Band 8A	Band 11	NDWI	AWEInsh	AWEIsh	LSWI	s	JM Distance	Number of selected features	Trend relative to previous test
15/07/2018		x	x	x	x	x	x	x	x	x	1,412	9	0,142%
18/09/2018	x	x	x	x	x	x	x	x	x		1,382	9	1,39%
10/10/2018	x		x	x	x	x	x	x	x	x	1,364	9	2,10%
3 dates		x	x	x	x	x	x	x	x	x	1,350	9	1,18%

TABLE C. 3 – Best subsets of features obtained for the different dates by letting the algorithm freely choose among the following reduced set of features (Band 2, Band 4, Band 8A, Band 11, NDWI, AWEInsh, AWEIsh, LSWI and s) to which the Band 8 has been added

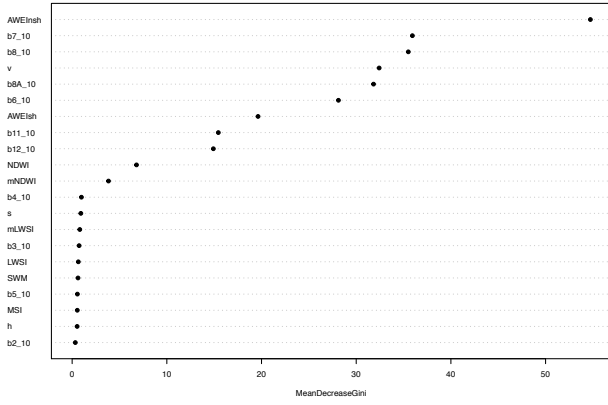
Dates	Band 2	Band 4	Band 8	Band 7	Band 8A	Band 11	NDWI	AWEInsh	AWEIsh	LSWI	s	JM Distance	Number of selected features	Trend relative to previous test
15/07/2018	x	x	x		x	x	x	x	x	x	x	1,412	10	=
18/09/2018	x	x	x		x	x	x	x	x	x	x	1,395	10	0,9%
10/10/2018	x	x	x		x	x	x	x	x	x	x	1,374	10	0,7%
3 dates	x	x	x	x		x	x	x	x	x	x	1,308	10	-3,11%

TABLE C. 4 – Best subsets of features obtained for the different dates by letting the algorithm freely choose among the following reduced set of features (Band 2, Band 4, Band 8A, Band 11, NDWI, AWEInsh, AWEIsh, LSWI and s) to which the Bands 8 and 7 have been added

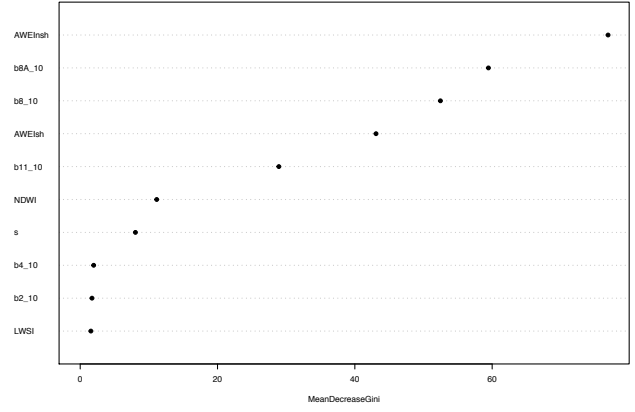
Dates	Band 2	Band 4	Band 8	Band 7	Band 8A	Band 11	NDWI	mNDWI	AWEInsh	AWEIsh	LSWI	s	JM Distance	Number of selected features	Trend relative to previous test
15/07/2018	x	x	x		x	x	x	x	x	x	x	x	1,413	11	0,07%
18/09/2018	x	x	x		x	x	x	x	x	x	x	x	1,402	11	0,5%
10/10/2018	x	x	x	x		x	x	x	x	x	x	x	1,392	11	1,31%
3 dates	x	x	x		x	x	x	x	x	x	x	x	1,384	11	5,81%

TABLE C. 5 – Best subsets of features obtained for the different dates by letting the algorithm freely choose among the following reduced set of features (Band 2, Band 4, Band 8A, Band 11, NDWI, AWEInsh, AWEIsh, LSWI and s) to which the Bands 8 and 7, as well as the indice mNDWI have been added

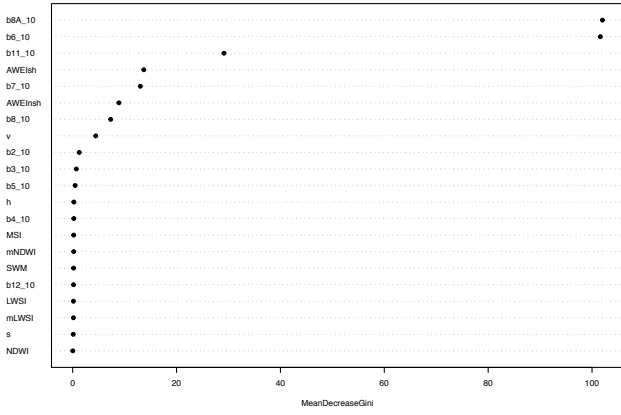
2. With Gini Index



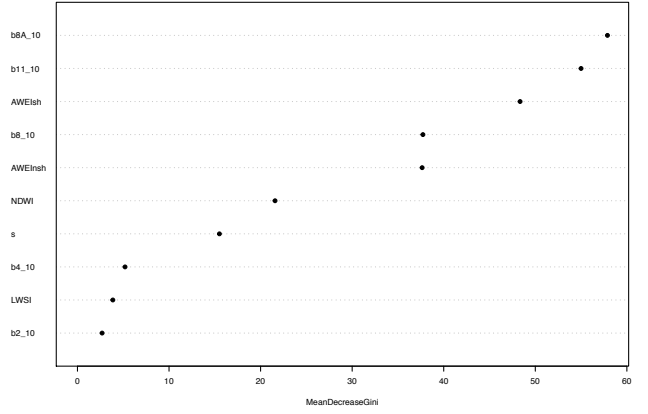
(a) 18/09/18 - Complete



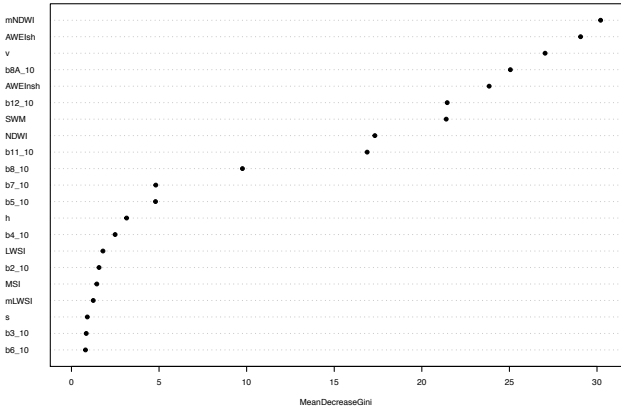
(b) 18/09/18 - Subset



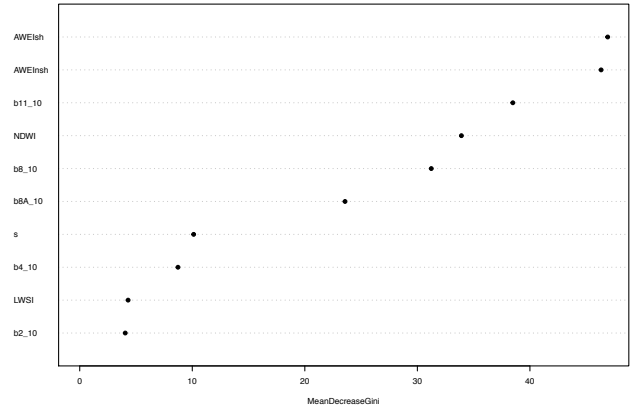
(c) 15/07/18 - Complete



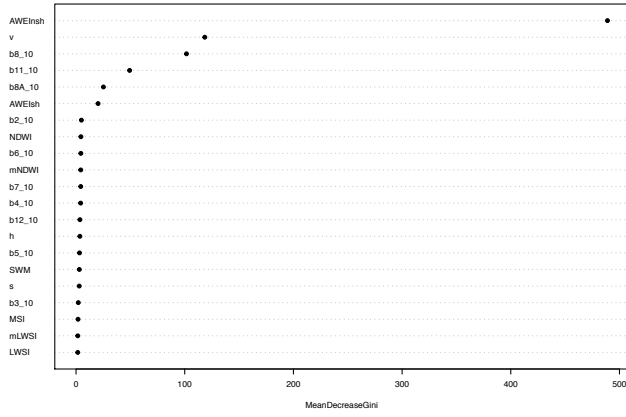
(d) 15/07/18 - Subset



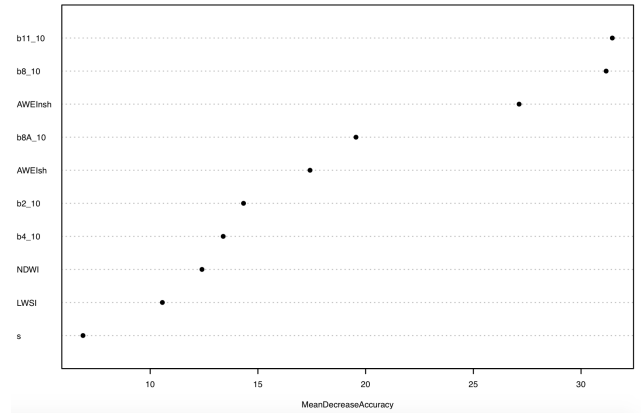
(e) 10/10/18 - Complete



(f) 10/10/18 - Subset



(g) 3 dates - Complete



(h) 3 dates - Subset

FIGURE C. 1 – Feature importance ranking depending on their decrease effect on the Gini index for discrimination between the categories : "Water-Geomembrane" and "Concrete-Vegetation". The ranking is illustrated for the eight RF classification models developed to calibrate the parameters of number of features and number of images.

Appendix D. Accuracy Assessment of the Classification Model

1. Calibration of the Number of Training Areas

TABLE D. 1 – Confusion matrices of the four RF classification models developed to calibrate the number of training areas. The calibration/validation image was restricted to the September image and the predicted categories were set up to : "Water-Geomembrane" and "Concrete-Vegetation".

(a) Situation 1 - Excellent SB				(b) Situation 2 - Homogeneous excellent SB				(c) Situation 3 - Homogeneous excellent SB + additional training areas				(d) Situation 4 - Only additional training areas representative of the categories : water-geomembrane and vegetation-concrete			
RF model 1		Classified		RF model 1		Classified		RF model 1		Classified		RF model 1		Classified	
Reference	W-G	1063	7	Reference	W-G	1025	45	Reference	W-G	930	140	Reference	W-G	683	387
	V-C	13	428		V-C	33	408		V-C	3	438		V-C	0	441
Precision		0.9879	0.9839	Precision		0.9688	0.9007	Precision		0.9968	0.7578	Precision		1	0.5326
Recall		0.9935	0.9705	Recall		0.9579	0.9252	Recall		0.8692	0.9932	Recall		0.6383	1
F1-Score		0.9907	0.9772	F1-Score		0.9633	0.9128	F1-Score		0.9286	0.8908	F1-Score		0.7792	0.6950
OA		0.9868		OA		0.9484		OA		0.9054		OA		0.7439	
CI		0.9796 - 0.9919		CI		0.9360-0.9590		CI		0.8895-0.9197		CI		0.7211-0.7657	

2. Calibration of the Number of Features and Images

TABLE D. 2 – Confusion matrices of the eight RF classification models developed to calibrate the number of features and number of images. Training areas were restricted to homogeneous excellent SB and the predicted categories were set up to : "Water-Geomembrane" and "Concrete-Vegetation".

(a) 18/09/18 - Complete				(b) 18/09/18 - Subset				(c) 15/07/18 - Complete				(d) 15/07/18 - Subset			
RF model 1		Classified		RF model 1		Classified		RF model 1		Classified		RF model 1		Classified	
Reference	W-G	30	3	Reference	W-G	30	3	Reference	W-G	14	19	Reference	W-G	20	19
	V-C	2	219		V-C	2	219		V-C	6	215		V-C	0	215
Precision		0.9375	0.9865	Precision		0.9375	0.9865	Precision		0.7000	0.9188	Precision		1	0.9208
Recall		0.9091	0.9910	Recall		0.9091	0.9910	Recall		0.4242	0.9728	Recall		0.4242	1
F1-Score		0.9231	0.9887	F1-Score		0.9231	0.9887	F1-Score		0.5283	0.9450	F1-Score		0.5957	0.9587
OA		0.9803		OA		0.9803		OA		0.9016		OA		0.9252	
CI		0.9547-0.9936		CI		0.9547-0.9936		CI		0.8581-0.9353		CI		0.8856-0.9544	

(e) 10/10/18 - Complete				(f) 10/10/18 - Subset				(g) 3 dates - Complete				(h) 3 dates - Subset			
RF model 1		Classified		RF model 1		Classified		RF model 1		Classified		RF model 1		Classified	
Reference	W-G	12	0	Reference	W-G	12	0	Reference	W-G	71	28	Reference	W-G	71	28
	V-C	10	245		V-C	9	246		V-C	35	628		V-C	33	630
Precision		0.5454	1	Precision		0.5714	1.0000	Precision		0.6698	0.9573	Precision		0.6827	0.9574
Recall		1	0.9608	Recall		1.0000	0.9647	Recall		0.7172	0.9472	Recall		0.7172	0.9502
F1-Score		0.7058	0.9636	F1-Score		0.7272	0.9820	F1-Score		0.6926	0.9522	F1-Score		0.6995	0.9538
OA		0.9625		OA		0.9663		OA		0.9173		OA		0.9199	
CI		0.9322- 0.9819		CI		0.937-0.9845		CI		0.8955 - 0.9359		CI		0.8984-0.9382	

TABLE D. 3 – Comparison of classification accuracy of the three calibration/validation images (multi-date) between model trained with the September calibration/validation images (mono-date) and with trained with the three calibration/validation images (multi-date) respectively. Training areas were restricted to homogeneous excellent SB and the predicted categories were set up to : "Water-Geomembrane" and "Concrete-Vegetation".

(a) Model trained with the September calibration/validation image (b) Model trained with the three calibration/validation images

RF model 1		Classified	
		W-G	V-C
Reference	W-G	3024	197
	V-C	103	1199
Precision		0.9671	0.8589
Recall		0.9388	0.9209
F1-Score		0.9527	0.8888
OA		0.9337	
CI		0.9260 - 0.9408	

RF model 1		Classified	
		W-G	V-C
Reference	W-G	3037	184
	V-C	121	1181
Precision		0.9617	0.8652
Recall		0.9429	0.9071
F1-Score		0.9522	0.8856
OA		0.9326	
CI		0.9249 - 0.9397	

3. Validity of the Model for Medium and Excellent with Reference Images

TABLE D. 4 – Accuracy obtained when applying the retained model (Trained with 24 homogeneous SB training areas on the September image with a subset of features to predict 2 categories) on the three reference images for excellent and medium SB, with traditional and modal comparison.

(a) Retained model on homogeneous Excellent SB (b) Retained model on homogeneous Excellent SB-Mode (c) Retained model on heterogeneous Excellent SB (d) Retained model on heterogeneous Excellent SB-Mode

RF model 1		Classified	
		W-G	V-C
Reference	W-G	2427	57
	V-C	78	1182
Precision		0.9689	0.9548
Recall		0.9774	0.9381
F1-Score		0.9731	0.9463
OA		0.9642	
CI		0.9577-0.9699	

RF model 1		Classified	
		W-G	V-C
Reference	W-G	31	4
	V-C	1	35
Precision		0.9688	0.8974
Recall		0.8857	0.9722
F1-Score		0.9254	0.9333
OA		0.9296	
CI		0.8433-0.9767	

RF model 1		Classified	
		W-G	V-C
Reference	W-G	598	140
	V-C	25	17
Precision		0.9599	0.1083
Recall		0.8103	0.4048
F1-Score		0.8788	0.1788
OA		0.7885	
CI		0.7581-0.8166	

RF model 1		Classified	
		W-G	V-C
Reference	W-G	11	8
	V-C	2	0
Precision		0.8462	0
Recall		0.5789	0
F1-Score		0.6874	0
OA		0.5238	
CI		0.2978-0.7429	

(e) Retained model on Medium SB (f) Retained model on Medium SB - Mode

RF model 1		Classified	
		W-G	V-C
Reference	W-G	208	168
	V-C	105	503
Precision		0.6645	0.7496
Recall		0.5532	0.8273
F1-Score		0.6037	0.7865
OA		0.7226	
CI		0.6934-0.7503	

RF model 1		Classified	
		W-G	V-C
Reference	W-G	22	44
	V-C	19	73
Precision		0.5366	0.6239
Recall		0.3333	0.7935
F1-Score		0.4119	0.6985
OA		0.6013	
CI		0.5204-0.6782	

Appendix E. Accuracy Assessment of the Temporal Evolution of Water Proportion for Medium and Excellent SB

1. Illustration of water proportion's evolution

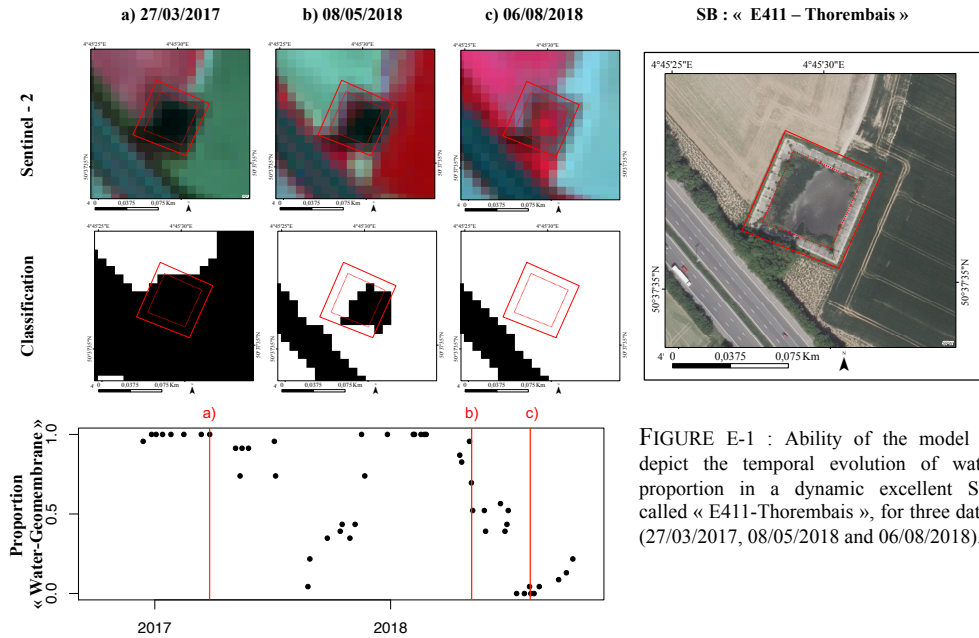


FIGURE E-1 : Ability of the model to depict the temporal evolution of water proportion in a dynamic excellent SB, called « E411-Thorembais », for three dates (27/03/2017, 08/05/2018 and 06/08/2018).

2. Algae Bloom

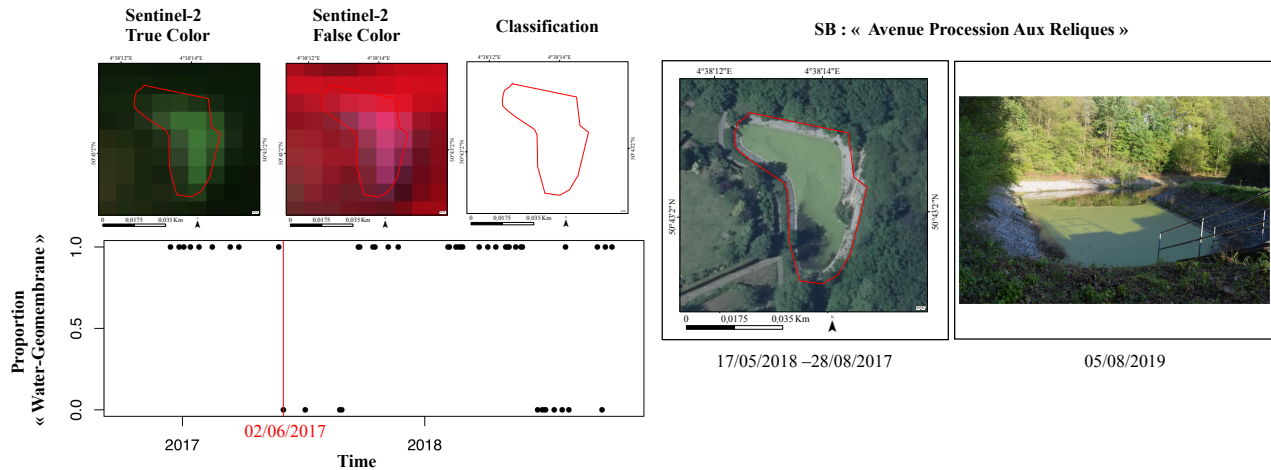


FIGURE E. 2 – Illustration of the occasional underestimation of water proportion caused by the presence of algae at the SB surface.

3. Validation Results

Date of the image	Correctly classified	Misclassified				Accuracy
		Landscape Feature Shadow	Lightness aberration	Algae	Other	
27/12/2017	2	0	83	0	0	0,02353
03/01/2017	0	0	84	0	0	0,0000
27/03/2017	60	13	1	0	0	0,7229
16/05/2017	62	1	0	0	9	0,9841
07/07/2017	75	0	0	2	1	0,9740
29/08/2017	75	6	0	2	2	0,8824
22/02/2018	59	3	0	0	20	0,7195
18/04/2018	69	9	0	0	0	0,8846
02/07/2018	79	0	0	2	1	0,9634
30/09/2018	66	15	1	0	0	0,8049

TABLE E. 1 – Validation performed on 10 images for all the SB covered by these images

ID	Name of the SB	Correctly classified	Misclassified				Accuracy
			Landscape Feature Shadow	Lightness aberration	Algae	Other	
9	Les Forges	31	12	2	0	3	0,6458
71	Bosquet	43	3	4	0	1	0,8431
122	Etang des roseraies	42	0	3	0	1	0,9130
7	Nodebais	26	15	8	0	1	0,5200
37	Heytens	34	1	4	0	5	0,7727
19	E411 - Thorembais	31	0	8	0	9	0,6458
21	E411 - Corroy-le-Grand	39	0	3	0	0	0,9286
67	Sablère	42	0	4	0	0	0,9130
84	Rue Emile Macq	33	0	4	0	7	0,7500
76	Chemin Saint thiry	42	0	5	0	0	0,8936

TABLE E. 2 – Validation performed on 10 SB (Excellent) for all the available images

ID	Name of the SB	Correctly classified	Misclassified				Accuracy
			Landscape Feature Shadow	Lightness aberration	Algae	Other	
30	Avenue de Mérode	34	0	7	0	2	0,7907
109	Bois Brûlé	41	1	2	0	0	0,7885
133	Corroy E411	45	0	3	0	1	0,9184
22	Le Ry Ternel	30	8	2	0	3	0,2308
68	Jaudrinia	37	0	2	0	0	0,7551
39	Rue de Pallant	42	0	5	0	0	0,8936
49	Bouston RER* (Depuis sa creation)	30	7	4	1	7	0,5660
80	BO Golf du Bosquet	39	0	2	0	0	0,8864
161	BO 27	30	7	4	0	11	0,5769
132	Chemin de la Cocquièrre	21	14	5	0	0	0,5250

TABLE E. 3 – Validation performed on 10 SB (Medium) for all the available images

Monitoring of stormwater basins with Sentinel-2 image time series over the province of Walloon Brabant (Belgium).

Master's thesis presented by Caroline Vassart

Scientific and public attention for flood risk management are constantly growing in the recent years because flash floods have the propensity to affect populations and societies, sometimes leading to dramatic humanitarian crises. It is therefore necessary to ensure an integrated and coordinated management of flooding mitigation measures.

However, the inventory and monitoring methods of stormwater basins (SB) that are currently used by the Walloon administration do not provide a comprehensive understanding of the location and the operational state of SB at both regional and provincial levels.

Remote sensing constitutes in that regard an unexploited and promising method for acquiring that information at those geographical scales. In fact, the recent launch of ESA's Sentinel-2 satellites has brought new perspectives for any development of an improved regional and provincial scale SB monitoring.

The following research study builds on that renewed perspective while at the same time offers a continuity to the earlier initial project of data centralization launched by the Walloon Brabant Province. The study aimed at assessing the feasibility of monitoring the hydrological functioning of SB using Sentinel-2 images time series over the province of Walloon Brabant.

This goal has been met through three steps. The first one consisted in the delineation of SB in order to constitute the SB dataset and the characterization of their potential detectability according to the adequacy between its characteristics and the limitations of Sentinel-2 images. Through the second step, a Random Forest (RF) classification model was developed to detect the presence of water inside the observable SB based on the spectral information provided by three Sentinel-2 calibration/validation images. The last step was dedicated to the application of the validated classification model on the time series of Sentinel-2 images in order to extract the temporal evolution of water inside potentially observable SB.

This study demonstrated that a majority (140/156) of SB, initially recorded in the provincial platform of Walloon Brabant, were observable and delineated by means of orthophotoplans. The assessment of SB detectability suggested that a majority (85/140) of SB was potentially observable from Sentinel-2 images. These results have further strengthened the confidence in the fact that the Sentinel-2 MSI sensor seems to be highly adequate for SB monitoring. Furthermore, the implemented methodology makes possible to determine *a priori* whether Sentinel-2 would be able to observe the SB and thus monitor its water surface.

This study concluded that Sentinel-2 image time series provides a cost-free mean for acquiring information about the elementary state and the temporal evolution of water presence in the observable SB over the province of Walloon Brabant. In fact, the successfully developed RF classification model was able to represent the temporal evolution of water with an accuracy of 72%. The 50 cloud-free observations spread over the years 2017 and 2018 provided a competitive observation frequency in comparison with visual inspection techniques.

Future research should explore the possibility of integrating Sentinel-1 images in the proposed approach in order to address the cloud obscuration and the artifacts caused by heavy precipitation and wet snow.

Key words : Remote Sensing, Stormwater Basins, Sentinel-2 Time Series, Detectability, Monitoring.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

Faculté des bioingénieurs

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