urlib.net/www/2022/02.08.01.22-TDI

REFERENCE SAMPLE SELECTION FOR SUPERVISED CLASSIFICATION OF A LOWER RESOLUTION IMAGE AIDED BY A HIGHER RESOLUTION IMAGE

Sabrina Paes Leme Passos Corrêa

Master's Dissertation of the Graduate Course in Remote Sensing, guided by Drs. Luciano Vieira Dutra, and Derek D'Arcy Lichti, approved in December 16, 2021.

URL of the original document:
<http://urlib.net/ QABCDSTQQW/46B5A2S>

INPE
São José dos Campos
2021

PUBLISHED BY:

Instituto Nacional de Pesquisas Espaciais - INPE
Coordenação de Ensino, Pesquisa e Extensão (COEPE)
Divisão de Biblioteca (DIBIB)
CEP 12.227-010
São José dos Campos - SP - Brasil
Tel.:(012) 3208-6923/7348
E-mail: pubtc@inpe.br

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES

PÁTRIA AMADA BRASIL GOVERNO FEDERAL

urlib.net/www/2022/02.08.01.22-TDI

REFERENCE SAMPLE SELECTION FOR SUPERVISED CLASSIFICATION OF A LOWER RESOLUTION IMAGE AIDED BY A HIGHER RESOLUTION IMAGE

Sabrina Paes Leme Passos Corrêa

Master's Dissertation of the Graduate Course in Remote Sensing, guided by Drs. Luciano Vieira Dutra, and Derek D'Arcy Lichti, approved in December 16, 2021.

URL of the original document:
<http://urlib.net/ QABCDSTQQW/46B5A2S>

INPE
São José dos Campos
2021

Cataloging in Publication Data

Corrêa, Sabrina Paes Leme Passos.

C817r Reference Sample Selection for supervised classification of a lower resolution image aided by a higher resolution image / Sabrina Paes Leme Passos Corrêa. – São José dos Campos : INPE, 2021.

xxviii + 201 p. ; (urlib.net/www/2022/02.08.01.22-TDI)

Dissertation (Master in Remote Sensing) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2021.

Guiding : Drs. Luciano Vieira Dutra, and Derek D'Arcy Lichti.

1. Remote sensing. 2. Supervised image classification. 3. Reference data selection. 4. Reference data quality. 5. Spatial data quality. I.Title.

CDU 528.584

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS
Serviço de Pós-Graduação - SEPGR

DEFESA FINAL DE DISSERTAÇÃO DE SABRINA PAES LEME PASSOS CORRÊA BANCA Nº317/2021, REG 430777/2019

No dia 16 de dezembro de 2021, as 13h30min, por teleconferência, o(a) aluno(a) mencionado(a) acima defendeu seu trabalho final (apresentação oral seguida de arguição) perante uma Banca Examinadora, cujos membros estão listados abaixo. O(A) aluno(a) foi APROVADO(A) pela Banca Examinadora, por unanimidade, em cumprimento ao requisito exigido para obtenção do Título de Mestra em Sensoriamento Remoto. O trabalho precisa da incorporação das correções sugeridas pela Banca e revisão final pelo(s) orientador (es).

Título: "Reference Sample Selection for supervised classification of a lower resolution image aided by a higher resolution image".

Observações da banca: O orientador é do exterior e vou encaminhar a declaração membro estrangeiro com as considerações dele (anexa ao processo).

Membros da Banca:

Dr. Thales Sehn Körting - Presidente - INPE
Dr. Luciano Vieira Dutra - Orientador - INPE
Dr. Derek D'Arcy Lichti - Orientador - Department of Geomatics Engineering - University of Calgary
Dr. Carlos Alberto Felgueiras - Membro Interno - INPE
Dra. Marinalva Dias Soares - Membro Externo - EMBRAER
Dr. Raian Varga Maretto - Membro Externo - Faculty of Geo-Information Science and Earth Observation, University of Twente

Declaração de aprovação do Dr. Derek D'Arcy Lichti anexa (8941442)

---

sei! assinatura eletrônica — Documento assinado eletronicamente por **Thales Sehn Korting**, **Pesquisador**, em 03/02/2022, às 17:20 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

---

sei! assinatura eletrônica — Documento assinado eletronicamente por **Carlos Alberto Felgueiras**, **Tecnologista**, em 04/02/2022, às 10:03 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

---

sei! assinatura eletrônica — Documento assinado eletronicamente por **Raian Vargas maretto (E)**, **Usuário Externo**, em 08/02/2022, às 14:10 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

---

sei! assinatura eletrônica — Documento assinado eletronicamente por **LUCIANO VIEIRA DUTRA (E)**, **Usuário Externo**, em 08/02/2022, às 22:40 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Marinalva dias soares (E)**, **Usuário Externo**, em 10/02/2022, às 21:37 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site http://sei.mctic.gov.br/verifica.html, informando o código verificador **9377753** e o código CRC **C9779B90**.

Referência: Processo nº 01340.008661/2021-28 SEI nº 9377753

"We cannot go back and start over, but we can begin now and make a new ending".

James R. Sherman
in *"Rejection"*, 1982

*To all those who have suffered from **mental illness** during the COVID-19 pandemic. And all the **neurodivergents** around the world. You can do it.*

ACKNOWLEDGEMENTS

There are a lot of people that I am incredibly grateful for. Not only the ones who were part of this masters degree journey, but specially the ones who gave me the opportunities to do so. To start, I am grateful for having Dona Thais as my mother; she has believed in me since my first blink of eyes. She has been there for me in all moments and taught me the most important lesson one could ever know: love is above all.

I also reminded all teachers and professors I have had. They are responsible for great part of this manuscript. Specially, I am grateful for my former professors back in Australia, Dr. Petra Helmholz and Dr. David Belton, who have first believed in me as a researcher and gave me the first opportunities to do so. I am also grateful for my forever professor, Dr. Afonso de Paula dos Santos, a great friend, someone I can count on and someone who has been there for me, someone who has believed that I could go anywhere.

And during my masters degree process, I am thankful for all my professors mainly Dr. Elizabete Caria Moraes and Dr. Sidnei João Siqueira Sant'Anna, they have been good friends and were always happy to help. I am truly grateful for having Dr. Luciano Vieira Dutra as my supervisor, for and believing in me even in the darkest moments and being there with me teaching and teaching how to be a researcher; for always taking me out of my comfort zone and making me think outside the box. You, professor, have such a messy and genius mind! I am also thankful for my co-supervisor Dr. Derek D. Lichti, someone who has all my admiration and who has inspired me to keep going. To you, my professors, you have my eternal gratitude.

And writing this manuscript, by far, has not been an easy task. For that, I am thankful for all support and recommendations from the board: Dr. Thales Sehn Körting, Dr. Carlos Felgueiras, Dr. Marinalva Dias Soares and my dear friend Dr. Raian Vargas Maretto. Despite all the pandemic situation, you have somehow helped me, either lecturing, having conversations, or just listening.

Still during the masters journey, I am thankful for the ABPG and all those who I am lucky to have met! Dr. Antonio Bertachini, Simone Del Ducca, Maria Tereza and Dr. Rafael dos Santos are people who I express my admiration and gratitude. Also, all members and former members of ABPG: you are the best! Well, I cannot forget the friends I made from ANPG, people who opened up my mind to see the world more differently. Also, how could I forget the *Bolinho Caipira de Jaleco* (BCJ) and my beloved forever friends from this group to spread science around Brazil, mainly Maíra Terra!

There is not way one can do a masters degree by oneself. And my friends reminded me constantly me that I was not alone during this entire process. They have been with me through my darkest moments and when I was most scared. This I would never forget and I appreciate you all for that! The "Dutrados", including Mariane Reis who have heard me and helped me enormously. Also ASA Room 56, the guys I have spent a lot of time with and I really miss you in a daily basis. And, well, SER-2019!! You guys are wonderful! And I could not forget the other people from INPE that have been there for me: Carol Cairo, Rogério Flores Jr., Luciano Ritter, Marlucia. I could never forget my dearests from Viçosa, Thales Heck, Gabriel Bhering and the ones from Camilo Chaves and ACEAK. Also my friends from Cabo Frio: Lucas De Lima, Alexya da Silva Pinto, Ana Beathriz Almeida de Souza and Ricardo Vargas, my dear friends, I wish you guys the best! Lastly but not least, how could I forget the person who has helped the most: Daniela Graner.

Finally, I would like to express my gratitude to Brazilian National Institute for Space Research (INPE), and specially the Earth Observation and Geoinformatics Division, for all the infrastructure and intellectual support it has given me. Also, this research was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001, and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), for which I am grateful.

ABSTRACT

One crucial step in Remote Sensing analyses is the collection of reference samples,used either to train supervised classifiers or to assess the accuracy of results. Up to date, it is common for reference samples to be treated as the truth, mainly due to the lack of feasible ways to assess their quality. However, the increased availability of very high spatial resolution imagery may create new opportunities for such a task. This thesis aimed to study the effect of the quality of reference data in supervised image classification. For this, a methodology was proposed to define and assess the quality of reference data, using a higher-resolution image as an auxiliary data. A controlled situation was used, with the high-resolution image, first sampled (by average) to a much lower-resolution image (LR) and then using LR to be the working classification. In this way it was guaranteed that there were no geometric errors. Besides, reference pixel quality was defined in terms of the relative proportion of the land cover class inside the area of the LR pixel. The same set of classes were used to classify the high-resolution image and histograms of classes inside the LR pixel were evaluated. The class of the LR pixel is then assigned to the mode of this class' histogram, provided that the mode is greater than 50% of total potential count inside the LR pixel. Then, sets of reference data, composed by those that met this minimum relative proportion criteria (i.e, $\geq$ 50%), were divided into training and test data using six different quality combinations (Setups) to be analysed. This study observed that the quality of reference data is a relevant factor for remote sensing analysis and for their respective spatial data quality; also, that the quality of test samples tends to be more influential to qualitative measures (thematic accuracy and *kappa* index) than the of the classifier itself. Finally, for this specific scenario, completeness analyses for classification accuracy, producer and user accuracy (related to omission and commission errors, respectively) played a more important role to represent the quality of an image than a thematic accuracy analyses.

Keywords: Remote Sensing. Supervised image classification. Reference data selection. Reference data quality. Spatial Data Quality. Classification Accuracy Assessment. Digital Image Processing. Machine Learning. Pattern Recognition.

SELEÇÃO DE AMOSTRAS DE REFERÊNCIA PARA CLASSIFICAÇÃO SUPERVISIONADA DE UMA IMAGEM DE MENOR RESOLUÇÃO AUXILIADA POR UMA IMAGEM DE MAIOR RESOLUÇÃO

RESUMO

Um ponto trivial em análises em Sensoriamento Remoto é a coleção de amostras de referência usadas tanto para treinar classificadores supervisionados, quanto para avaliar a qualidade dos resultados da classificação. Até o momento, é comum tratar amostras de referência como verdade, principalmente devido a falta de formas práticas para avaliar essa qualidade. No entanto, o aumento da disponibilidade de imagens de altíssima resolução espacial pode criar novas oportunidades para esta tarefa. A partir disto, este documento teve como objetivo estudar o efeito da qualidade em dados de referência para classificação supervisionada de imagens. Para isso, uma metodologia foi proposta para definir a qualidade de dados de referência utilizando uma imagem de resolução mais alta como dado auxiliar. Uma situação controlada foi utilizada, onde a imagem de mais alta resolução foi reamostrada (por média) para uma imagem de mais baixa resolução (LR) de forma que LR foi usada como imagem a ser trabalhada nas análises. Desta forma, foi garantido que não houvesse erros geométricos de registro. Além disso, a qualidade de pixels de referência foi definida em termos de proporção relativa da classe de cobertura de terra dentro da área de um pixel LR. O mesmo grupo de classes de cobertura foi utilizado para classificar a imagem de mais alta resolução para se avaliar os histogramas das classes dentro de um pixel LR. Daí, as classes dentro de um pixel LR foram assinaladas para a moda do histograma daquela classe considerando que a moda tem frequência maior que 50% do total de pixels dentro de um pixel LR. Então, os conjuntos de dados de referência compostos por dados que tenham atingido o critério de mínima proporção relativa (i.e, $\geq 50\%$) foram divididos entre dados de treino e teste utilizando seis combinações de qualidade de dados de referência (Setups) para ser depois analisados. A partir disto, este estudo observou que a qualidade de dados de referência é um fator relevante em análises de sensoriamento remoto e para seu controle de dados cartográficos; ainda, que a qualidade de dados de teste tende a ser mais influente para medidas de acurácia temática (acurácia global e índice *kappa*) que o tipo de classificador utilizado. Por fim, para este cenário específico, elementos de completude, acurácia de produtor e do usuário (relacionadas aos erros de omissão e comissão, respectivamente) se mostraram peças mais importantes para representar a qualidade de uma imagem do que análises temática.

Palavras-chave: Sensoriamento remoto. Classificação supervisionada de imagens. Seleção de dados de referência. Qualidade de dados de referência. Controle de Qualidade Cartográfica. Avaliação da Qualidade da Classificação. Processamento digital de imagens. Aprendizado de máquina. Reconhecimento de padrões.

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

DT	–	Decision Trees
DT5	–	Decision Trees C5.0
ESA	–	European Spatial Agency
ET-CQDG	–	Technical Specification - Spatial Data Quality
GeoDMA	–	Geographic Data Mining Analyst
GIS	–	Geographic Information System
HR	–	Higher-resolution baseline image
ICA	–	International Cartographic Association
INDE	–	Brazil National Infrastructure of Spatial Data
INPE	–	Brazil National Institute for Space Research
ISO	–	International Organisation for Standardisation
KNN	–	K-Nearest Neighbours
KNN-5	–	K-Nearest Neighbours with K=5
LR	–	Lower-resolution image
MSI	–	MultiSpectral Imager
NIR	–	Near-infrared
OA	–	Overall accuracy
OOB	–	Out-Of-bag
PIX	–	Pixel-based Baseline classification
prop	–	Modal class proportion
QGIS	–	Quantum GIS
RBF	–	Radial basis function
REG	–	Region-growing Baseline classification
RF	–	Random Forest
RSS	–	Reference Sample Selection
SIRGAS 2000	–	Geocentric Reference System for South America
std	–	Standard deviation
SVM	–	Support Vector Machine
SVM-OAO	–	Support Vector Machine One-Against-One
SWIR	–	Shortwave infrared
TCI	–	True Colour Image
UTM	–	Universal Transverse Mercator
VIS	–	Visible

CONTENTS

Page

1 INTRODUCTION

Remote sensing, as defined by Campbell and Wynne (2011), is the usage of electromagnetic radiation in regions of the electromagnetic spectrum reflected or emitted from the Earth surface in order to derive information from it. Such information can be used for monitoring the Earth's surface: for military purposes (HUDSON; HUDSON,), land and maritime monitoring, as well as emergency management (ESA, 2015b).

Remote sensing imagery studies started with the launch of the first satellite[1] in 1957 and since then, there are several remote sensing satellites orbiting the Earth, with their resolutions improving alongside with science and technology (CAMPBELL; WYNNE, 2011). Therefore these breakthroughs aid scientists to have new understanding of environmental and socioeconomic dynamics of the Earth, which can be exemplified by deforestation as well as inland and ocean water monitoring (ELMES et al., 2020).

The evolution in science and technology has led to advances in the satellite sensors that have increased their resolutions with time. Regarding the spatial resolution, in this year of 2021, the availability of moderate resolution sensors ($5 - 30m$), high-resolution sensors ($1 - 5m$) and very-high resolution sensors ($< 1m$) has enabled more focused studies regarding Machine Learning techniques for image classification (ELMES et al., 2020).

According to Richards and Xiuping (2006), image classification can be defined as a computer-based quantitative analysis of the attributes of each pixel - which can be spectral bands available or their derivatives - so we can label them identifying as belonging to a particular set of pixels of interest. Plus, Lary et al. (2018) state that "machine learning is an automated approach to building empirical models from the data *alone*".

However, supervised machine learning (ML) algorithms are dependent on a "comprehensive representative set of examples", referred to as training data (LARY et al., 2018) as well as test (or validation) data (BRANDT; MATHER, 2009), which combined are entitled here as reference data. As these reference data are trivial for defining the predictive ML model as well as for assessing the classification quality, their quality seems to be important. In this sense, we come up with the question: *to what extent the quality of reference data affects the predictive model of a supervised image clas-*

[1]The first launched satellites was Sputnik-1 in 1957 by USSR, it had military positional purposes (LAUNIUS et al., 2002).

sification? In this study, we define the pixel quality as being related to its spectral purity, thus a pure pixel is considered if it is 100% pure (SHIMABUKURO; PONZONI, 2019).

There are some studies regarding this subject, though focusing on the quality of training data, such as Elmes et al. (2020) who organised a survey in this matter proving the importance of these data into machine learning algorithms. Foody et al. (2016) also studied the effect of imperfect reference training data and found that the errors related to training data depend on the involved classes. Another study shown in Mellor et al. (2015) focused in training data imbalance for machine learning prediction model. Another study was conducted by Jin et al. (2014) which assessed how training data selection affect binary image classification. There are also studies in which regard the sampling designs and image classification results (STEHMAN; CZAPLEWSKI, 1998; STEHMAN, 1999; STEHMAN, 2009; STEHMAN; FOODY, 2019).

We emphasise the existence of quantitative studies on reference data, such as Shimabukuro and Ponzoni (2019) who determine the quality of reference data based on *in situ* data using the so-called spectral mixture. However, these studies are entirely focused on applications, such as inland water monitoring (MACIEL et al., 2019; CAIRO et al., 2020), rather then the quality assessment of image classifications. Despite these studies present relevant conclusions, most of them are thematic accuracy rather than completeness elements, which flags a lack a completeness elements' studies regarding effect of training data in supervised remotely sensed image classification.

Another point to be addressed is a statement that have been taken as a convention in Pattern Recognition studies: for classifications, it is important to select test samples as representative as possible. Nonetheless, this fact can be understood as true for assessing the quality of the classifier *per se*, by analysing how well the classifier can allocate information to a certain label, which usually is discussed in Pattern Recognition books (RICHARDS; XIUPING, 2006; HASTIE et al., 2009; THEODORIDIS; KOUTROUMBAS, 2009). The quality of a thematic map, which is the resulting classified image, may not always be related to the quality of the classifier, specially for region-based classifiers, once the training object is different from the testing object; this idea emerged from Foody et al. (2016) where the involved classes influence the effect of training data quality.

In a nutshell, some points regarding the assessment of reference data for image classification are:

- Most studies focus on the assessment of training data;
- There is a convention that the quality of testing (validation) data for image classification must be as high as possible;
- The studies regarding the selection of reference data concern mostly the isage of thematic accuracy rather than completeness;
- So far, studies that define the quality of reference data depend on *in situ* information.

From these considerations and assuming that the acquisition of reference samples for supervised image classification is prone to the existence of sprectrally mixed pixels, in this study, we seek to answer the following question: *to what extent choosing more (or less) class-representative training and test samples affect the accuracy of image classification?* In order to answer that question, this study presented the objectives hereafter.

1.1 Objectives

For answering the question stated above, the main objective of this study is to control the quality of acquired remote sensing reference data for supervised image classification and then determine how it affects the thematic map accuracy. From this, the specific objectives are:

a) Propose a semi-automated method for selecting reference pixels with diverse mixture profiles aided by an adequately higher spatial resolution image, entitled Reference Sample Selection (RSS);

b) Assess the accuracy of image classification using thematic accuracy measures[2] (overall accuracy and *kappa* index) as well as completeness elements (commission and omission errors) varying the quality of training and test samples as measure of purity.

1.2 Manuscript structure

Therefore, this manuscript firstly presents a Literature Review (Chapter 2) of image classification, classifiers, reference data selection and classification accuracy assessment. Then, the Methodology is separated into two chapters: (i) Proposed Approach,

[2]Please refer to DSG (2016) for more information regarding the use of the expression "thematic accuracy" and "completeness" for classification accuracy.

shown in Chapter 3, where the theoretical concept of Reference Sample Selection (RSS) is presented alongside with the manual approach for selecting reference samples and (ii) Experimental Planning (Chapter 4), where the used material is stated as well as how RSS is applied to a controlled situation. After presenting the methodology, the Results and Discussions regarding two distinguished controlled situations are presented in Chapter 5. Finally, Conclusions and final remarks are shown in Chapter 6.

Additionally, aiming the accessibility of remotely sensed studies, most of the images in this study are colourblind friendly and the ones which could not meet this criteria are presented in Black and White format in Appendix C.

2 LITERATURE REVIEW

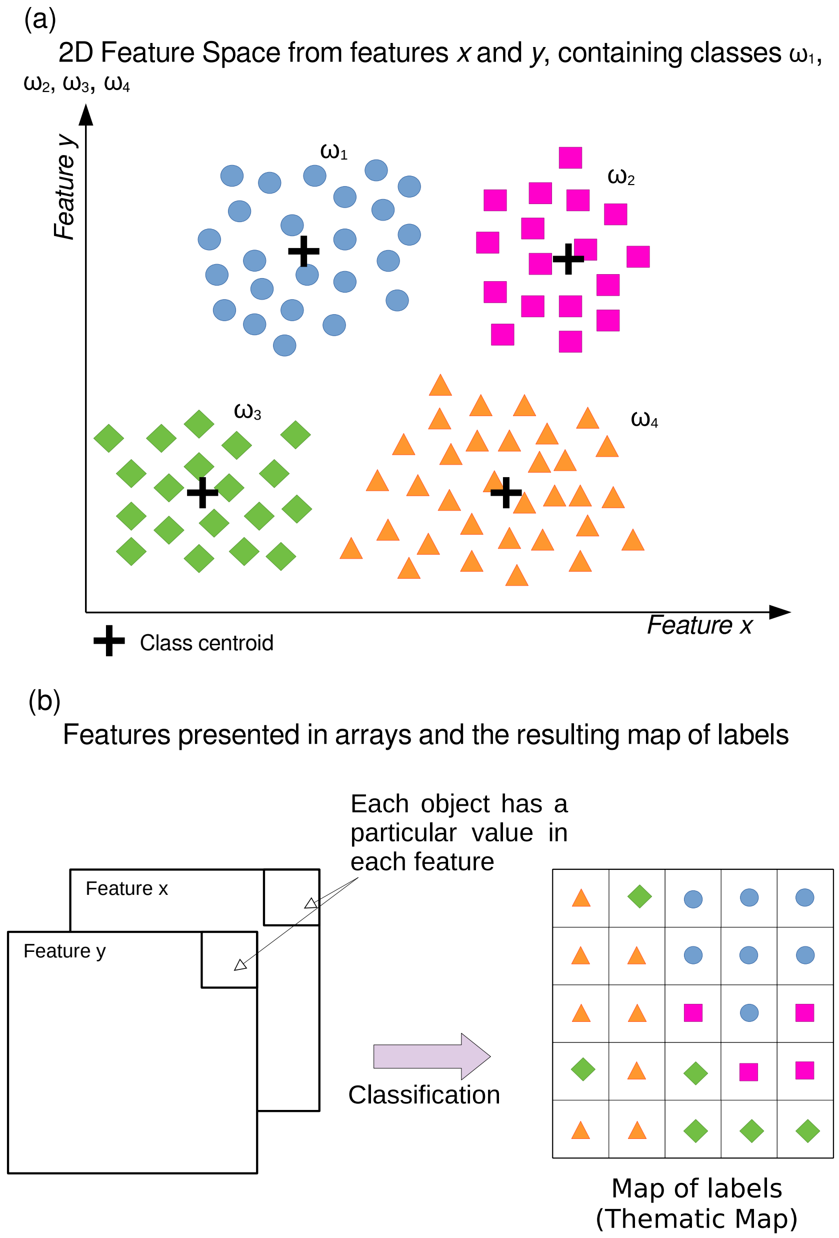
2.1 Classification systems

Classification is a quantitative analysis as it analyses the image numeric properties, as stated by Richards and Xiuping (2006) and Novo (2010) and it is effective for a general assessment of ground cover types according to its geometric characteristics. This analysis can be summarised in assigning a meaning to, or labelling, an object of the image. The object of the image can be a pixel or a set of pixels, so-called region. When the object is a pixel, then the classification is entitled herein as Pixel-based Classification discussed in Section 2.2.1. When the object is a set of pixels, also known as regions, and metrics derived from this set of pixels is computed, there is a segmentation followed by a classification based on these metrics, entitled herein as Region-based Classification, discussed in Section 2.2.2.

Moreover, according to Brandt and Mather (2009), the image classification uses a set of input features in a k-dimensional space, so-called feature space, so it can create a relationship between any pattern and a labelled class or ground cover type. For a remotely sensed data, these features may be values of spectral reflectance, metrics derived from these reflectances or even geographical information, such as slope and elevation, as stated by Brandt and Mather (2009). These features can be selected or extracted, i.e, derived from selected features which are chosen according to their importance to the classification model as stated by Theodoridis and Koutroumbas (2009).

For a visual understanding of the feature space, Figure 2.1(a) presents a 2-dimensional feature space where the features are x and y. Each i^{th} object in the feature space has its features values, being (x_i, y_i) as a set of coordinates vector. According to these coordinates in the feature space, they can be classified into the classes ω_1, ω_2, ω_3 and ω_4. Once they are classified, the results are presented in a map of labels or Thematic Map, as showed in Figure 2.1(b) which is an array representation regarding their array coordinates. Nonetheless, Figure 2.1 shows the classes well-separable and it also presents the classes centroids for a better understanding of the classes.

Figure 2.1 - Illustration of classification using two features.



Classification represented in two forms: (a) in the feature space and (b) in array as a Thematic Map.

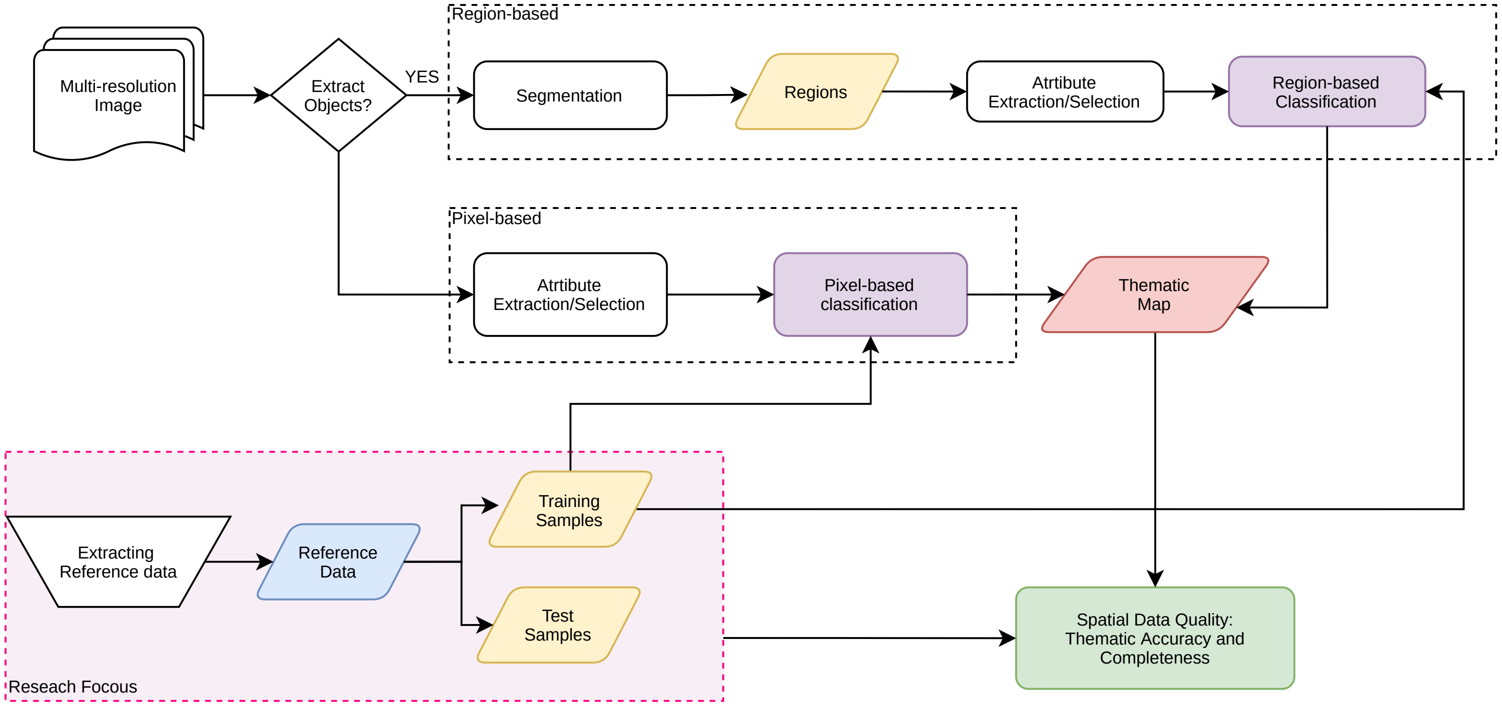
SOURCE: Adapted from Richards and Xiuping (2006) and Theodoridis and Koutroumbas (2009).

Once the feature space in defined, the labelling, or classification approach takes place. It can be supervised, unsupervised or a combination of both (called semi-supervised or weakly supervised), according to Theodoridis and Koutroumbas (2009) and Brandt and Mather (2009). To be more specific, Richards and Xiuping (2006) described unsupervised classification as classifying objects according to feature classes without previous knowledge of the ground cover classes. On the other hand, the supervised approach needs a user to collect data or samples to train the classifier, so it can delineate the class region in the feature space (BRANDT; MATHER, 2009). This study focuses on supervised classification, hence all subsequent topics will imply this approach.

In a nutshell, Figure 2.2 presents a step-by-step flowchart of a supervised image classification, based on Richards and Xiuping (2006), Brandt and Mather (2009), and Theodoridis and Koutroumbas (2009). This figure illustrates both region-based and pixel-based classification scenarios. Regarding the former, it is necessary the step of segmentation to form regions and later the extraction and selection of attributes prior to the classification. On the other hand, the latter has only attribute extraction and selection prior to the classification. Nonetheless, both cases need the extraction of reference data generating training and test samples, which is the main focus of this study. This extraction is thoroughly discussed in Section 2.2.

The next sections explain in more details the Pixel-Based and Region-based classifications for remotely sensed data.

Figure 2.2 - Flowchart of usual supervised classification system for remotely sensed data.



This flowchart summarises a supervised classification procedure for region-based and pixel-based classifications. It also emphasises this study focous.

SOURCE: Author.

2.2 Reference data selection for supervised classification

Selecting reference data is the process of selecting pixels or regions that characterise the thematic classes of interest, as defined by Brandt and Mather (2009). These data is further used by pattern recognition algorithms in Earth Observation studies as well as for classification accuracy assessment. Ideally, the reference data should concern the entire region of study; however, as this is unpractical, selecting these samples using statistics becomes necessary (STEHMAN, 2009). The author defines sample as "a subset or portion of the region mapped" hence these samples should summarise the data.

Considering they should summarise the data, the quality of these data must be put into thought. There are two main sources of remotely sensed reference data: those selected from basemaps and *in situ* samples. The former as they are easier to acquire, they are the majority of the cases and, as stated by Elmes et al. (2020), must consider characteristics in which may influence the selection judgement, such as sun angle, spectral band selection and image contrast. The latter regards the collection of samples directly in the field which are usually defined as points. In both scenarios, the samples are called "ground truth" (STEHMAN, 2009; ELMES et al., 2020) regardless of its reliability.

Besides, in remote sensing image classification, these samples are objects which can be pixels or regions, as stated by Richards and Xiuping (2006). According to the authors, the first case, the pixels selected have their feature or terrain values, such as grey scale, digital number directly measured or any derivative from these metrics; the latter, a metric of the region is computed, such as mean, median or mode.

The error source of acquiring reference data is presented in Section 2.3. Then, the statistics considerations required for selecting samples, or reference data, in this study will be called sampling design and are discussed in Section 2.3.1 and studies concerning sampling size are exposed in Section 2.3.2. The accuracy assessment is dealt separately in Section 2.5.

2.2.1 Pixel-based classification

The Pixel-based classification is, for remotely sensed data, the most used approach because it infers directly on the pixel data for each used feature(BRANDT; MATHER, 2009). For pattern recognition, Richards and Xiuping (2006) states the pattern as a pixel vector containing the i^{th} pixel values for each band in the format $(x_i, y_i, ..., n_i)$.

The features can be values of brightness, spectral reflectance, digital level, slope, elevation or aspect, or metrics derived from these features (RICHARDS; XIUPING, 2006; BRANDT; MATHER, 2009). Metric derived from features are defined as feature extraction, which is when the feature is a derivation two or more other features. As an example, the bands red and near-infrared (NIR) that together generate the normalised difference vegetation index (NDVI), used for better recognition of vegetation in optical satellite images (BRANDT; MATHER, 2009; NOVO, 2010).

2.2.2 Region-based classification

As stated by Brandt and Mather (2009), it is possible to create a set of pixels and classify the set instead of using the pixels separately. Blaschke (2010), Hossain and Chen (2019) define segmentation as partitioning an image in a set of different regions according to specific properties, such as shape, texture, colour and digital level. These regions are segments, entitling this procedure as segmentation. According to Baatz et al. (2008), this is a two-stage approach where the first process segments the image in sets of pixels and metrics and computed from it (KÖRTHING, 2012; ABDOLLAHI; PRADHAN, 2021); the second step is image classification and analysis. An illustration of an image segmentation is presented in Figure 2.3.

As a difference from the Pixel-based, Espindola et al. (2006) and Baatz et al. (2008) affirm that this approach has more advantages, creating a more consistent thematic map. The authors also mention that segmentation requires some input parameters, which affect the quality of the outcomes. Besides, Baatz et al. (2008) highlight that the created objects impact the classification results, once "the accuracy and the significance of the final measurements, numbers, and statistics directly and actually critically depend on the quality of the segmentation".

The segmentation technique used herein in the Multiresolution segmentation that, according to Blaschke (2010), it was first proposed in Baatz and Schäpe (2000) and later developed into programmable workflow in Baatz et al. (2008).According to the authors, this process used Cognition Network Technology (CNT), are based on hierarchical networks of objects. In this sense, it is possible to address different object classes with distinct object modification strategies. This algorithm can processed in the software *eCognition 9.1* (Trimble Germany GmbH, 2014) and they set some input parameters[1] to be used:

- Image Layer weights: the weight of any used band in the image;

[1]We point out that this algorithm is a blackbox, therefore it cannot be well described.

Figure 2.3 - Illustration of an image segmentation.

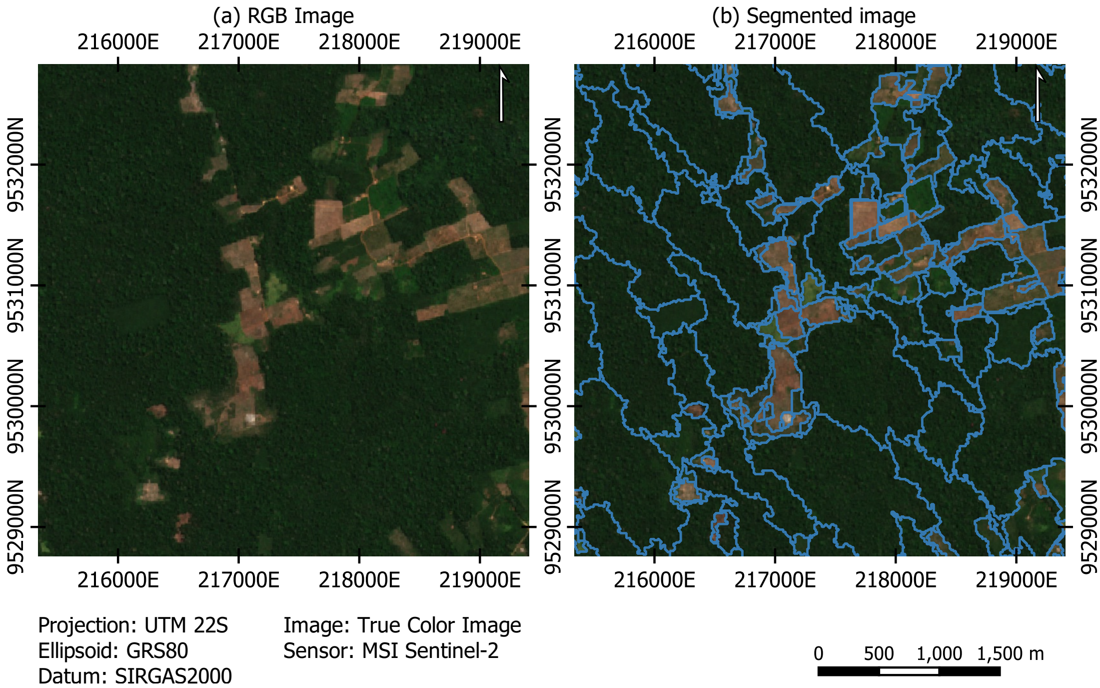


Image in RGB composition. True Color Image (TCI) band from Sentinel-2 MSI Sensor taken in 9^{th} August 2020.

SOURCE: Author.

- Scale parameter: the scale of the object, where a higher scale tends to create larger segments;
- Shape: defined the weight that the shape criterion would have in the process;
- Compactness: the weight of compactness criterion.

In a second moment, features are extracted from the objects to be selected as reference data. The *eCognition 9.1* (Trimble Germany GmbH, 2014) can extract several features regarding the object shape (e.g. area, asymmetry, perimeter and border length), spectral features (mean, mode and standard deviation). The extraction can use Grey Level Co-occurrence Matrix (GLCM) or Gray-Level Difference Vector (GLDV), proposed by Haralick et al. (1973). The image classification works in these features as region vectors, for example, the j^{th} region is represented by the vector (A_j, s_j, m_j), the same format as the pixel vectors.

2.3 Error source on classification systems

According to Elmes et al. (2020), acquiring training reference data[2] for pattern recognition if often done by the same approaches and usually are subjected to the same errors despite the fact that not necessarily they are relevant for assessing the impact of errors in the reference data.

These approaches regard how the reference data is obtained concerning (i) its purity, i.e. how this sample characterises the studied label, whereas it comes from another sensor or not, or if the sample relates to ground measurements; (ii) the sample design, which, as explained by Stehman and Foody (2019) is the protocol of how the reference data will be selected for later assessment. Both are to be reported herein.

Moreover, Elmes et al. (2020) also state few problems associated with training reference data, being: (i) the demand for reference data due to the great amount of data; (ii) most of the reference data relies on human-generated products which may inherit errors; (iii) the uncertainty of reference data acquisition is barely reported or assessed, usually are considered as perfectly accurate and (iv) error propagation from the not satisfactory reference data quality to further data analysis.

2.3.1 How to acquire reference data: sampling design

Brandt and Mather (2009) defines the sampling design, or as they call, sample scheme, as selecting objects that characterise the thematic classes of interests from a certain population. Elmes et al. (2020) state that the sampling design refers to "where, when, how many, and what type of samples are placed". Besides that, Warner et al. (2009) and Olofsson et al. (2014) affirm that the sampling design forms the basis of the thematic map accuracy assessment and it must be done regarding the specific objectives of the accuracy assessment.

Moreover, Stehman (1999) defines seven major sampling design criteria discussed by Stehman (2009), in which aids the analyst to determine the optimum sampling design for their need, being:

- **C1** - *The sampling protocol satisfies the requirements of a probability sampling design*: it justifies the accuracy estimates which are derived from the sample;

[2]As mentioned in the Introduction, studies analysing accuracy of reference data are most commonly focused in training data rather than training and test data.

- **C2** - *The sampling design must be practical*: is it realistic to expect that the protocol will be implemented correctly? This criteria considers the limitations of collecting and analysing the reference data properly;

- **C3** - *The design must be cost effective*: is the funding sufficient for the sampling design?

- **C4** - *The sample is spatially well-distributed*: spatially distributed samples have a tendency of being more precise than those without a good spatial distribution;

- **C5** - *The sampling variability of the accuracy estimates should be small:* the obtained accuracy estimate should be close to the same value independently of the sample selected in a way the the assessment would be repeatable;

- **C6** - *Sampling variability or precision of the accuracy estimator should be estimated without undue reliance on approximations other than those related to sample size*: as some sampling designs require approximate standard errors;

- **C7** - *Ability to accommodate a change in sample size at any step in the implementation of the design*: this considers the fact the accuracy assessment budgets may be unpredictable and may change after the protocol has been initiated.

These sampling designs criteria can aid to determine the sampling design to use for a specific study (STEHMAN, 2009) and, it depends on what are the requirements of the pattern recognition approach is being used in case of training samples (ELMES et al., 2020).

Stehman (1999), Stehman (2009),Olofsson et al. (2014), Stehman and Foody (2019) and Elmes et al. (2020) present four sampling schemes: simple random, stratified random, clustering and systematic. These designs can be combined forming other sampling schemes depending on the study objective. We present here Simple Random Sampling as well as Stratified Sampling. For more information, please refer to Cochran (1977) and Stehman (1999).

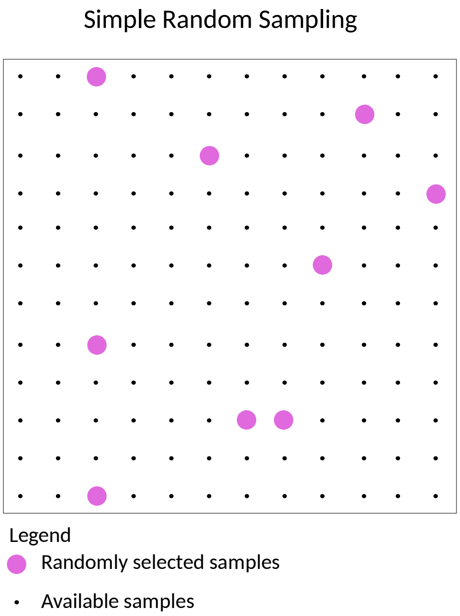
2.3.1.1 Simple random sampling

Simple random sampling is the idea of selecting samples randomly, as its name define it. As it is random, it infers a probability sampling design (C1), as defined by Stehman (2009) and it also infers simplicity (C2) as defined by Stehman (1999). This sampling design is also easily adapted to further changes in the sample size (C7) as defined by the author.

Usually, this design is used with another sampling design. According to Stehman (2009), Olofsson et al. (2014), this sampling design can be directly applied to select a set of clusters, a two-stage cluster sampling, which is a set of objects within a cluster, or a set of objects within a stratum.

To illustrate this sampling design, Figure 2.4 presents randomly selected sample from a box. We highlight that this sampling design alone may not meet spatial distribution condition (C4).

Figure 2.4 - Sampling Design: Simple Random Sampling example.



SOURCE: Author.

2.3.1.2 Sampling design: stratified random sampling

When they are divided into classes, the number of samples for each stratum can be defined so a precise estimate can be ensured (OLOFSSON et al., 2014). Each stratum size determination have a different objective (STEHMAN, 2009; OLOFSSON et al., 2014; STEHMAN; FOODY, 2019).

Stratified random sampling is the design recommended by Olofsson et al. (2014) as good practice. According to Stehman (2009) strata are groups of objects that belong to one stratum in a way that the strata form part of the entire object population. The author highlights that, for remote sensing, these strata can be divided according to the map classes or according to their spatial location.

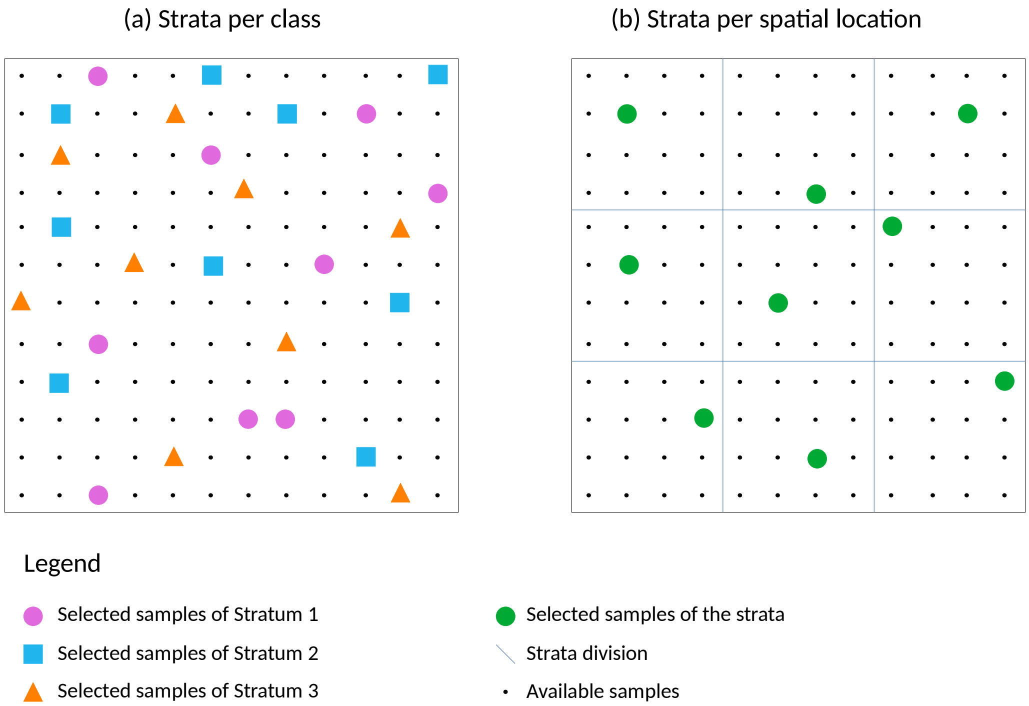
Figure 2.5 illustrate these two types of stratified sampling showing strata per class in Figure 2.5(a), where the samples are grouped into different map classes and strata per location, shown in Figure 2.5(b), where the samples are chosen according to their location within the cells, so samples spatial distribution is achieved (C4) (STEHMAN, 1999).

When they are divided into classes, the number of samples for each stratum can be defined so a precise estimate can be ensured (OLOFSSON et al., 2014). Obviously, the stratum size will depend on availability of resources to collect it, therefore this availability will be a constraint to determine it, as pointed out by Stehman and Foody (2019).

Each stratum size determination has a different objective (STEHMAN, 2009; OLOFSSON et al., 2014; STEHMAN; FOODY, 2019). When there is an equal allocation of sample size, all classes are considered to have equal importance, as mentioned by the authors. On the other hand, if there is a difference in importance, priority classes should have a greater sample size, so the confidence interval can be narrowed, as explained by Stehman and Foody (2019). According to Stehman (1999), this design may be used to certify that sample sizes for specific stratum meet the requirement for each stratum (C6).

As drawback, this design may not be reliable when there are subgroups, or strata, that are not of interest due to a poor distribution of sample among the strata which are now the aim of the accuracy assessment (STEHMAN, 1999). Therefore, the author emphasises that the best use of stratified sampling design when we are nearly sure to retain the importance of identified strata.

Figure 2.5 - Sampling Design: Stratified Sampling examples.



Strata from stratified sampling considering (a) strata per class, where the selected samples are in groups of classes and (b) satrata per location where the samples are selected within their regions divided with grey lines.

SOURCE: Adapted from Stehman (1999).

2.3.1.3 Comparing sampling designs

As Stehman (2009) studied specifically the sampling designs regarding the accuracy assessment, they also compared their relative strengths according to the design criteria, as presented in Table 2.1. We emphasise that the table presents all sampling designs from the paper aiming to compare the effectiveness of the stratified sampling with other sampling designs.

From Table 2.1, all designs where there is simple random protocol are considered as strong regarding flexibility criteria (C7). The author highlights that stratified (map class) random design is a prime candidate for class-specific accuracy objective, and it is the most commonly employed design. In agreement with that, Olofsson et al. (2014) recommends the use of stratified sampling as a good practice for remote sensing image classification.

Table 2.1 - Relative strengths ans weaknesses of basic sampling according to desirable design criteria.

Design	C1	C2	C3	C4	C5	C6	C7
Simple Random	●	●	○	○	○	●	●
Systematic	●	●	○	●	○	○	○
Stratified (map classes) random	●	●	○	○	●	●	●
Stratified (map classes) systematic	●		○		●	○	○
Stratified (spatial) random ($n_s = 1$)	●	●	○	●	○	○	
Stratified (spatial) random ($n_s > 1$)	●	●	○		○	●	●
Stratified (spatial) systematic	●	●	○	●	○	○	○
Cluster random	●		●	○	○	●	
Cluster systematic	●		●		○	○	○
Stratified random cluster	●		●	○			
Stratified systematic cluster	●		●			○	

The criteria are: C1) probability sample, C2) practical, C3) cost, C4) spatial balance, C5) precise estimates of class-specific accuracy, C6) ability to estimate standard errors, and C7) flexible to change in sample size. The rating symbols are • = strength and ○ = weakness; absence of a symbol indicates the design is 'neutral' with regard to that criterion. Besides, n_s means the number of selected samples in the stratum.

SOURCE: Stehman (2009).

2.3.2 Sample size

When it comes to sample size n, for both scenarios, training and test data, Brandt and Mather (2009) emphasise that nonparametric machine learning classifiers tend not to significantly affect the map accuracy when a small sample set is used. Additionally, Cochran (1977) mentions that larger sample set can result in waste of resources whereas too small sample sizes may bias the outcomes hence the sample size should be put into thought.

Nonetheless, Olofsson et al. (2014) and Cochran (1977) present methods for determining a minimum sample size according to the chosen sampling design. On the other hand, Brandt and Mather (2009) points out a rule of thumb of using a sample size 30 times the number of used features (bands)[1], which will be considered in this study.

[1]Considering this, the Brandt and Mather (2009) mention the case of hyperspectral data sets, where the sample size could be "unfeasibly large" hence they recommend the use of dimentionality reduction or feature selection techniques.

2.4 Machine learning classifiers

In machine learning classification, there are two main groups of classifiers: parametric and nonparametric, as mentioned by Richards and Xiuping (2006). The former assumes that the input data can be described using Gaussian probability distribution function in the feature space, which is generally the case for multispectral images, as the authors state. As example of parametric classifiers, Theodoridis and Koutroumbas (2009) mention Maximum Likelihood and Minimum Distance. The latter, non-statistical or nonparametric is *"an histogram approximation of a unknown probability distribution function"*, hence the data are not required to be Gaussian, as stated by Richards and Xiuping (2006). Some of the methodologies are Linear Discrimination, Support Vector Machine, Neural Networks, Decision Trees and Random Forests (RICHARDS; XIUPING, 2006).

This study is focused on nonparametric classification approaches and the ones used for this study are described herein. These approaches are K-Nearest Neighbours (KNN), Support Vector Machine (SVM), Decision Trees (DT) and Random Forests (RF).

2.4.1 K-Nearest Neighbours (KNN)

K-Nearest Neighbours (KNN) is one of the simplest approaches for image classification, being a nonparametric model and it does not require any assumptions regarding the input data (RICHARDS; XIUPING, 2006; ABDOLLAHI; PRADHAN, 2021). The authors also declare that its main hypothesis is that similar data tend to be neighbours in the feature space, hence it considers proximity between data for the predictions; in other words, it classifies the data based of k near neighbours samples. To define the nearest neighbours, few distances can be used and this a parameter for using this estimator (THEODORIDIS; KOUTROUMBAS, 2009). Among the distances, there is the Euclidean Distance (Equation 2.1) or the Minkowsky Distance (Equation 2.2). These distances are based on two samples, a and b regarding n features, hence, a pixel or region vector is defined by $a = (a_1, a_2, ..., a_{n-1}, a_n)$ is this example.

$$D_{E_{(a,b)}} = \sqrt{(a_1 - b_1)^2 + ... + (a_n - b_n)^2} = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2} \tag{2.1}$$

$$D_{M_{(a,b)}} = \left(\sum_{i=1}^{n} |a_i - b_i| \right)^{\frac{1}{r}} \tag{2.2}$$

Where,
$D_{E_{(a,b)}}$ is the Euclidean Distance;
$D_{M_{(a,b)}}$ is the Minkowsky Distance and
r is constant to be defined during the process.

As for choosing the k values, $k \in \mathbb{N}$ and it is usually an uneven value (ABDOLLAHI; PRADHAN, 2021). The lower the k value, more noises the classification may have. On the other hand, the higher the k value, more generalised the classification can be. In this sense, choosing k is an empiric process and it depends on the input data, although Abdollahi and Pradhan (2021) mention the use of some approaches.

Figure 2.6 illustrates how the KNN Classifier works in a 2D feature space, considering classes 1 (circles) and 2 (squares). The triangle is the object vector to be labelled. This figure uses $k = 1$, $k = 2$ and $k = 3$ neighbours and it shows the distance between the object vector in study to the training samples. When $k = 1$ and $k = 3$, the object vector is labelled as Class 1 (circle). However, when $k = 2$, the object vector cannot be labelled.

Figure 2.6 - Illustration of KNN classifier.

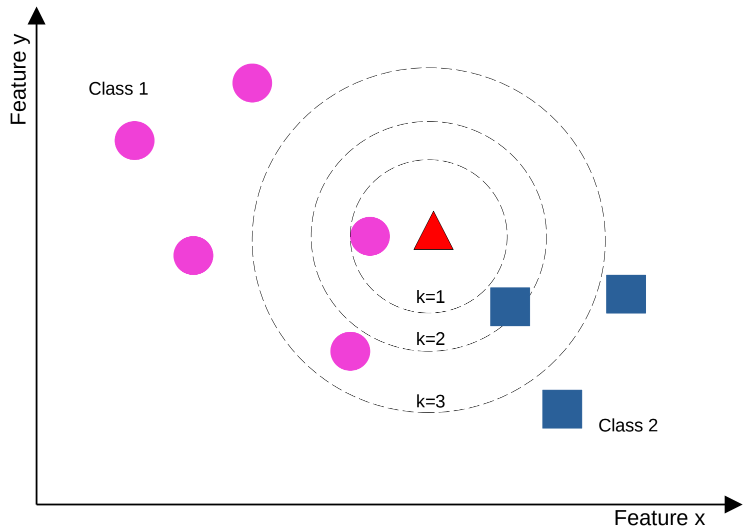


Illustration of the KNN Classifier, observing the red triangle sample in a 2D feature space. This sample is compared to Class 1 (pink circles) and to Class 2 (blue squares). When $k = 1$ and $k = 3$, the sample is allocated to Class 1 and when $k = 2$, the sample cannot be allocated.

SOURCE: Author.

Once the KNN model is fit, a decision boundary is created, as presented in Figure 2.7, which will determine the classification model. Moreover, Figure 2.7 also presents the classes centroid, which is the average value for all classified vectors and training samples allocated to that class. The centroid position express the importance of selecting samples as close to this centroid as possible for a more accurate supervised classification.

Additionally, Richards and Xiuping (2006) mentions that the nearest pixels can be weighted according to the inverse of the used distance. This leads us to two forms of performing the KNN Classifier: uniform, which considers the number of near pixels and weighted, which considers the inverse of the distance.

Figure 2.7 - Illustration of KNN classifier with decision boundary.

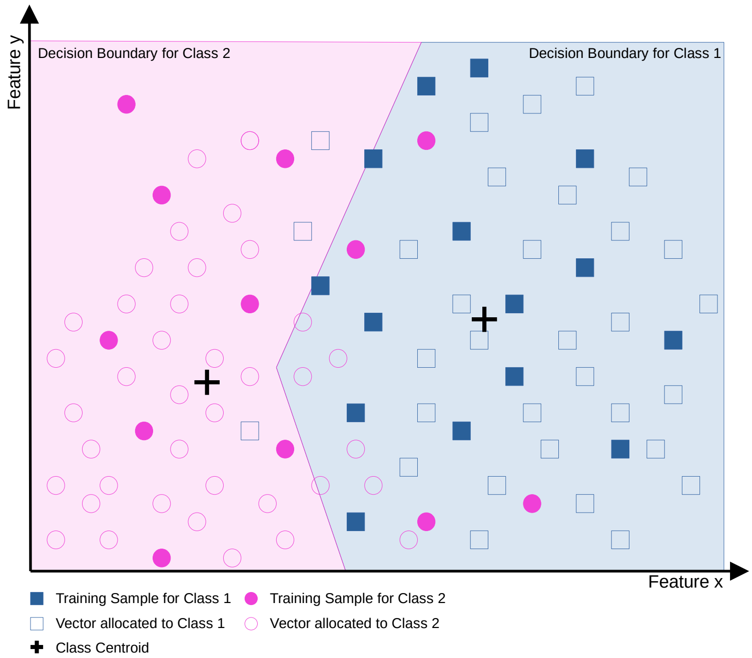


Illustration of the KNN Classifier, in a 2D feature space presenting the decision boundaries created. When selecting the training samples, it is important to select classes as close to the class centroid (crosses) as possible.

SOURCE: Author.

2.4.2 Support Vector Machine (SVM)

As the KNN classifier, the Support Vector Machine (SVM) is a nonparametric classifier in which is based on computing distances measures between data vectors to split the data according to decision boundaries in the feature space (CHANG; BAI, 2018). The difference is that SVM improves its robustness by creating a structural risk minimisation when filtering noises in the training process and that leads to a high classification accuracy, as stated by Brandt and Mather (2009), Chang and Bai (2018).

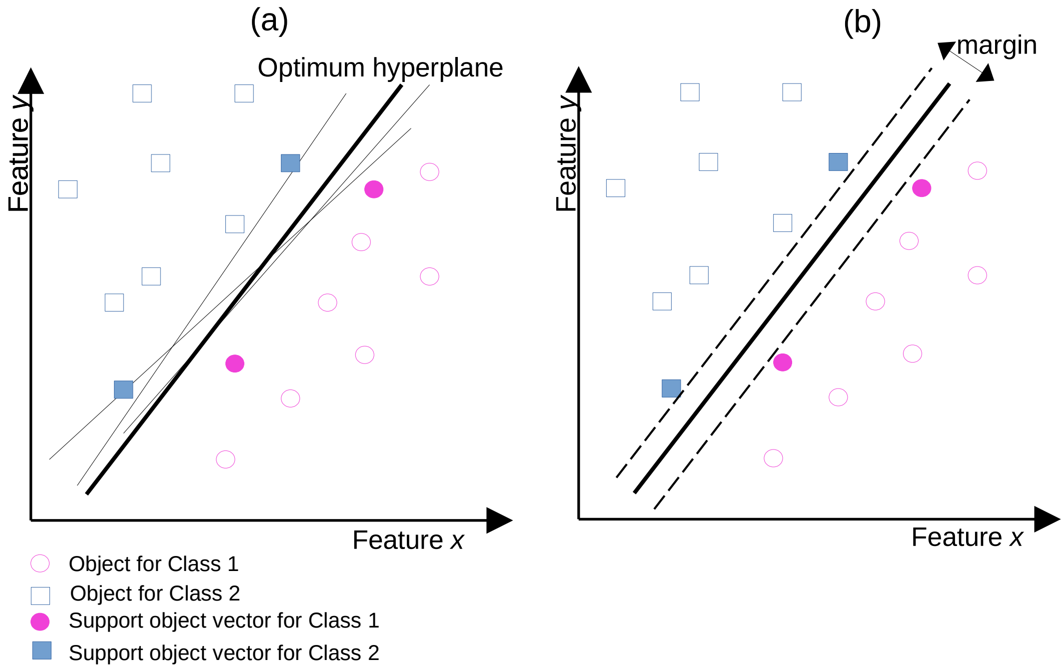
The decision boundary for the SVM classifier is a hyperplane, which constructing it is the main core of the classifier (BRANDT; MATHER, 2009). According to Richards and Xiuping (2006), the SVM uses only object vectors near the separating hyperplane, called support object vectors. Moreover, as stated by Theodoridis and Koutroumbas (2009), the hyperplane is not unique, hence there is a need to find an optimal hyperplane with the optimum distance from the support object vectors.

Likewise, Theodoridis and Koutroumbas (2009) state that an optimum hyperplane has the highest margin on both sides so there is a smaller risk of causing an error, hence being more trustworthy when dealing with unknown data. Another point regarding the hyperplane is that it can be n-dimensional, where n is the number of studied features (CHANG; BAI, 2018).

To illustrate how SVM works in a 2D feature space, Figure 2.8(a) presents few possibilities of linear hyperplanes and the selected optimum hyperplane with a thicker line. The optimum hyperplane is defined as the one with the highest margin from the near support object vectors, as showed in Figure 2.8(b).

When a linear hyperplane cannot separate the classes, it is named nonseparable and some non-linear approaches are used, called the non-linear SVM (BRANDT; MATHER, 2009). The authors elucidate that in this case, it is usual to change the number of used dimensions - usually for more - for a better separability of the hyperplane as showed in Figure 2.9. For nonlinear SVM, kernel functions are used where the most common ones are: linear kernel, polynomial kernel, radial basis function (RBF) and sigmoid function (CHANG; BAI, 2018).

Figure 2.8 - Illustration of SVM with the optimum hyperplane and the margin on a linear separable case.



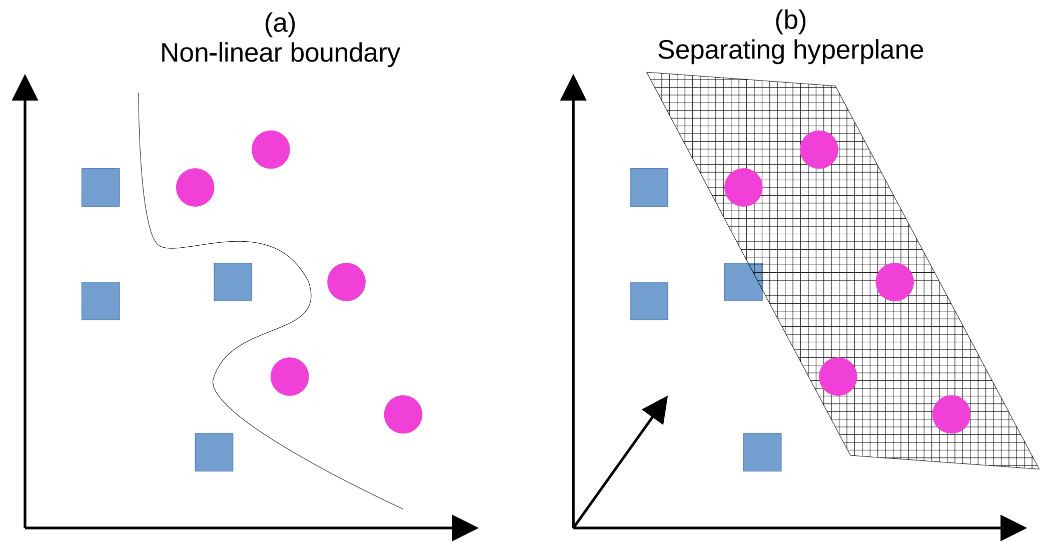
A 2D feature space illustrating (a) few possible hyperplanes according to the object support vectors and (b) the optimal hyperplane with the margin from the object support vectors. This Figure considers a linear separable case.

SOURCE: Adapted from Richards and Xiuping (2006), Brandt and Mather (2009) and Chang and Bai (2018).

To be more specific, concerning RBF, its equation is shown in Equation 2.3, showing how the feature space is split considering two features (x, y). For this function, γ defines the influence a training sample has in the modelling process; the other usual input parameter is C, defined how smooth the decision surface will be, i.e. the error penalty (BRANDT; MATHER, 2009).

$$K_{(x_i,y_i)} = e^{-\gamma\|x_i-y_i\|^2}, \gamma > 0 \tag{2.3}$$

Figure 2.9 - Illustration non-linear separation of the Support Vector Machine Classifier.



A (a) 2D feature space illustrating a non-linear boundary separating the classes and (b) a 3D feature space with the data mapped into a higher dimensional space to increase separability between classes.

SOURCE: Adapted from Brandt and Mather (2009).

All the explained above concerns a binary scenario, using only two classes. However, when there are more than two classes, there is the Multi-Class problem which consists of combining a set of binary classifiers so they can form a multi-class classifier (CHANG; BAI, 2018). The authors claim that there are two main approaches for doing so: one-against-one and one-against-all.

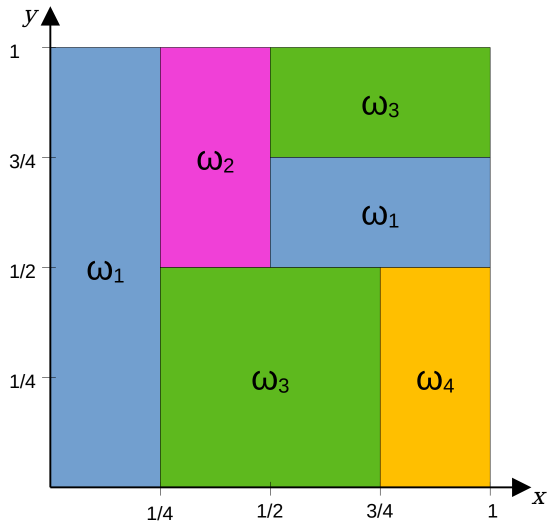
When it comes to one-against-one classification strategy, Brandt and Mather (2009) state that each classifier is trained on two of the M classes, with all possible combinations of one-against-one classes are evaluated from the M classes training set, being a total of $M(M-1)/2$ classes. For labelling an object, the majority of votes from the classifiers is used as a decision ruler (BRANDT; MATHER, 2009; CHANG; BAI, 2018).

Regarding the one-against-all strategy, there is M SVM classifiers for M classes and each classifier is trained to separate that class to the $M-1$ other classes, as explained by Brandt and Mather (2009) and Theodoridis and Koutroumbas (2009).

2.4.3 Decision Trees

Decision Trees (DT) is, as stated by Chang and Bai (2018), a nonparametric classifier known to be a rather efficient classifier, being capable of handling nonlinear relations between features and classes. It partitions the feature space into unique regions corresponding to the classes as illustrated in Figure 2.10.

Figure 2.10 - Decision Tree partition of a 2D feature space.



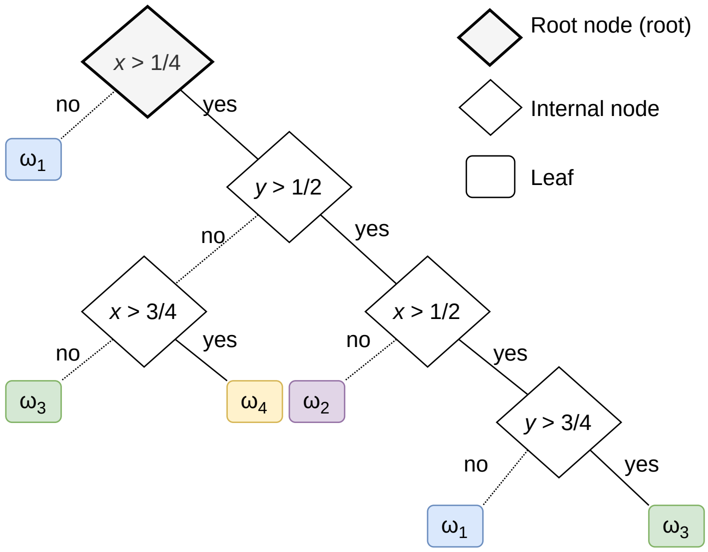
A 2D feature space (x, y) divided based on decision tree rule regarding four classes: ω_1, ω_2, ω_3 and ω_4.

SOURCE: Adapted from Theodoridis and Koutroumbas (2009).

The feature space in Figure 2.10 can be reinterpreted as presented in Figure 2.11, based on decisions regarding a certain feature in a tree known as ordinary binary classification trees (OBCTs), as stated by Theodoridis and Koutroumbas (2009), which consists of binary splits (yes, no) and each split has binary subsets.

The trees have nodes, and it starts with the root node which usually is the entire dataset and it is divided into internal nodes as stated by Chang and Bai (2018). What controls the tree growth is a stop-splitting rule, known as leaf, according to the Theodoridis and Koutroumbas (2009) and Chang and Bai (2018).

Figure 2.11 - Binary tree with decision nodes and leaves.



Binary decision tree regarding Figure 2.10 considering two features, x, y, and four classes ω_1, ω_2, ω_3 and ω_4.

SOURCE: Adapted from Theodoridis and Koutroumbas (2009).

Richards and Xiuping (2006) and Körthing (2012) explain that there are three tasks for designing the decision trees: finding the optimal tree structure, choosing the optimal subsets of features at each node and, finally, selecting the decision rule to use at each node; these tasks must be completed in a way that there is a minimum error rate, number of nodes or path length. In other words, the split rules are decided according to computations considering statistical probability rather than randomly (KÖRTHING, 2012).

When it comes to GeoDMA plugin, it uses a Decision Tree model entitled C5.0 (KÖRTHING et al., 2013). The Decision Tree C5.0 (DT5) algorithm is explained in Kuhn and Johnson (2013), which considers input parameters to optimise the splitting efficiency. These input parameters for the C5.0 decision tree classifier are presented in Table 2.2.

Table 2.2 - C5.0 Decision Trees input parameters for GeoDMA plugin.

Parameter	Description
Boosting	Boosting interactions or a single model to be used.
Trials	Number of boosting iterations.
Global Pruning	Using a global pruning to simplify the tree.
Minimum Cases	The lowest number of samples put into two or more splits.
Winnow	Option to make a feature selection prior to the classification.
Decomposed Rules	Whereas the tree is decomposed into a rule set based model.
Early Stopping	Whereas the internal method for stopping is used.
Subset	Whereas the model evaluates groups of splitting discrete predictions.
Fuzzy Thresholds	Evaluate possible advances data splits.
Confidence factor	The confidence factor for splitting the tree.

SOURCE: Körthing et al. (2013), Kuhn and Johnson (2013).

2.4.4 Random Forests

The Random Forest (RF) prediction model was developed by Breiman (2001) and its goal is to enhance the classification accuracy by using a combination of several decision trees, trained features and labelled data. According to Brandt and Mather (2009), each tree classifier is created by a random vector, which is sampled independently from the training set of the input vector data.

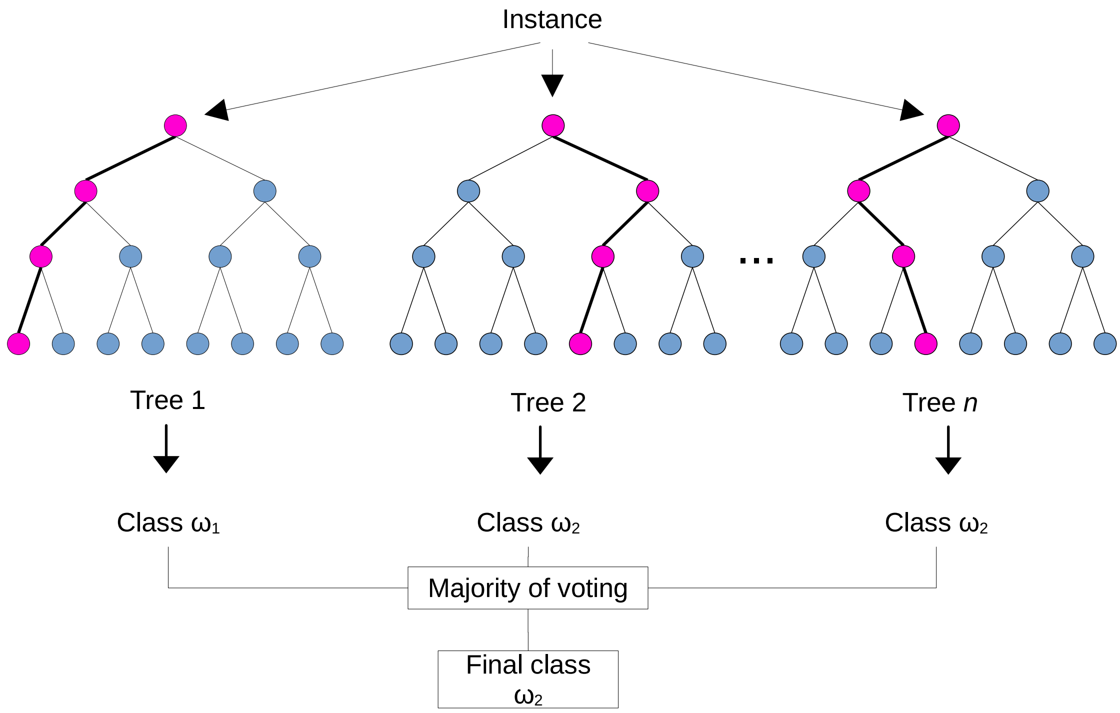
Each tree has a unique vote for the most popular class and the majority of the votes represent the assigned class (BRANDT; MATHER, 2009; HASTIE et al., 2009). To illustrate the random forest structure, Figure 2.12 presents an input instance with n random trees where most of the votes are to class ω_2, which is the final classification for that instance.

An important aspect of random forest classifier uses the idea of bagging for a random feature selection, where a feature split in each node is selected as the one with best outcomes among a set of n randomly chosen trees (THEODORIDIS; KOUTROUMBAS, 2009). The authors also state that this randomness characteristic effects virtually the performance improvement.

Additionally, as stated by Chang and Bai (2018), not all input samples are used as training; usually about 2/3 of it is randomly selected as training and the remaining third is used for a cross-validation to check the classifier performance. This performance, pointed out by Hastie et al. (2009), is almost identical to the one obtained using the k-fold cross-validation, hence statistically there may be no need for further validation[3].

This inner validation is entitled bagging, which, as described by Breiman (2001) can not only "give estimates of the generalisation error of the combined ensemble of trees" but it also "gives estimates for the classification strength and correlation" and the author entitles out-of-bag (OOB) estimates. The OOB is also considered as a feature selection once it computes the order of importance of the input features. Breiman (2001) also points out that the OOB estimate is unbiased.

Figure 2.12 - Random Forest structure.



Structure of random forest with n trees showing the majority of votes for class ω_2.

SOURCE: Adapted from Sothe (2019).

[3]This argument considers the quality of a classifier, although when it comes to a thematic map, i.e., classified remote sensing image, an external validation is mandatory according to international standards, ISO (2013) and the Brazilian regulation, DSG (2016).

2.5 Image classification accuracy assessment

When performing a supervised classification, training data must be used; however, it is trivial to assess the quality of the classification as stated by Richards and Xiuping (2006). Additionally, Brandt and Mather (2009) define training and test data as:

- *Train Data:* "are used in supervised pattern recognition to 'teach' a classifier the main characteristics of a class". They emphasise the need of a minimum sample size assigned to each class to ensure their proper representation;

- *Test Data:* are used for assessing the classification accuracy and are put aside during the image classification. When the thematic map is produced, these data are labelled using the same prediction model as the thematic map so a comparison between the predicted data and the 'true'[3] data is made.

Besides that, in this study we understand that *image classification accuracy can be different from classification accuracy* considering supervised classification depending on the unit object used. It is possible to train a prediction model using regions and assess it with pixels and *vice-versa* for remotely sensed data. In this case, there is a difference between them, as the classification accuracy requires the same type of object for training and test processes. This is not true for image classification accuracy, as we are assessing how the remotely sensed cartographic product is close to the "reality", considering land user and cover for that specific area. That said, we present literature studies for splitting the reference data when they are from the same type, i.e. when image classification accuracy is equal to the classifier's accuracy.

For splitting the reference data intro training and test samples, there are some statistical approaches that consider a finite number of samples and are presented in Section 2.5.1.

2.5.1 Splitting reference data into training and test samples

This is the first step for assessing the accuracy of a classifier and it is conducted prior to the classification depending on the splitting process, which may also be called as

[3]The usage of the expression "ground truth" is defined as misleading and inaccurate description as "it is human to err, and one presumes that this aspect of human behaviour extends to the collection of test and training data in remote sensing" (BRANDT; MATHER, 2009).

resampling. This section will present some approaches, such as resubstitution, cross-validation, bootstrap, leave-on-out and the holdout method (conventional split in remote sensing).

2.5.1.1 Resubstitution

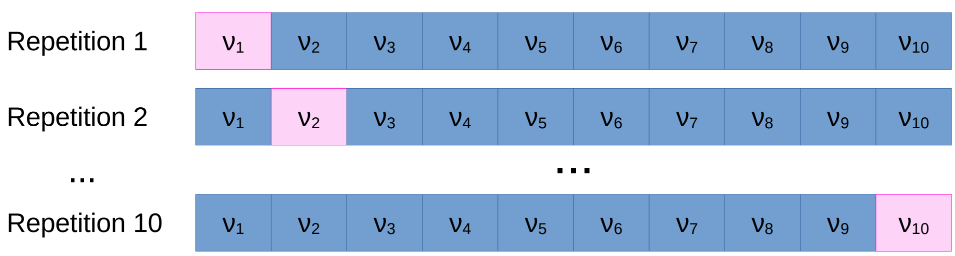
Theodoridis and Koutroumbas (2009) explains the resubstituion method as using the same dataset for both test and train data, which may bias the error probability optimistically.

2.5.1.2 *k*-fold Cross-validation

This is the recommended approach in the case of small reference data to be separated into training and test dataset (BRANDT; MATHER, 2009; HASTIE et al., 2009; JAMES et al., 2021). According to Hastie et al. (2009), Theodoridis and Koutroumbas (2009) and Lyons et al. (2018) in this approach, we divide the entire dataset in k approximately equal parts or subsets (ν), where the dataset is defined as $\{\nu_1, \nu_2, ..., \nu_k\}$. The first iteration, we holdout ν_1 as test and train the model with the set $\{\nu_2, ...\nu_k\}$. The second iteration, we holdout ν_2 and train the model with $\{\nu_1, \nu_3, ...\nu_k\}$ and so on as shown in Figure 2.13.

Figure 2.13 presents a scheme with $k = 10$ where lighter (pink) subsets are heldout as test data while blue subsets are training data. For each repetition a different subset is used as test data until all subsets are used as test data (BISHOP, 2006). Therefore, the model is always fit in $k - 1$ parts.

Figure 2.13 - k-fold Cross Validation scheme.



A scheme of the k-fold cross validation, with $k = 10$. It partitions the data into $k = 10$ subsets $\{\nu_1, ...\nu_{10}\}$. The test sample for each repetition is in pink while training samples are in blue.

SOURCE: Adapted from Bishop (2006).

The estimate for the k-fold cross validation for each repetition, as determined by Hastie et al. (2009) and James et al. (2021) is determined by the missclassified observations (Err). The CV error rate for the entire cross-validation is determined the same way.

$$CV_{(k)} = \frac{1}{k}\sum_{i=1}^{k} Err_i \tag{2.4}$$

Usually, the k-fold cross-validation uses $k = 5$ or $k = 10$ due to computational advantages and bias-variance trade off (JAMES et al., 2021).

2.5.1.3 Leave-One-Out cross-validation

When the number of folds in the k-fold cross validation is the same as the number of samples ($k = n$), then the technique name is *leave-one-out* cross-validation and it is recommended to be used when data is scarce (BISHOP, 2006). Additionally, Theodoridis and Koutroumbas (2009) claims that "the total number of errors leads to the estimation of the classification error probability" while keeping independence between training and test data.

According to Bishop (2006) and Hastie et al. (2009) this approach is rather expensive computationally due to the number of repetitions used.

2.5.1.4 Bootstrap

As stated by Theodoridis and Koutroumbas (2009) and Hastie et al. (2009), this approach is used when there is a limited amount of available data. The authors explain this method by using an initial sample S, with sample size n and resample it to S' with the same sample size n with repeated data as replacement. The test data will be the samples not used for training data. Each bootstrap sample is selected via simple random sampling from the original data.

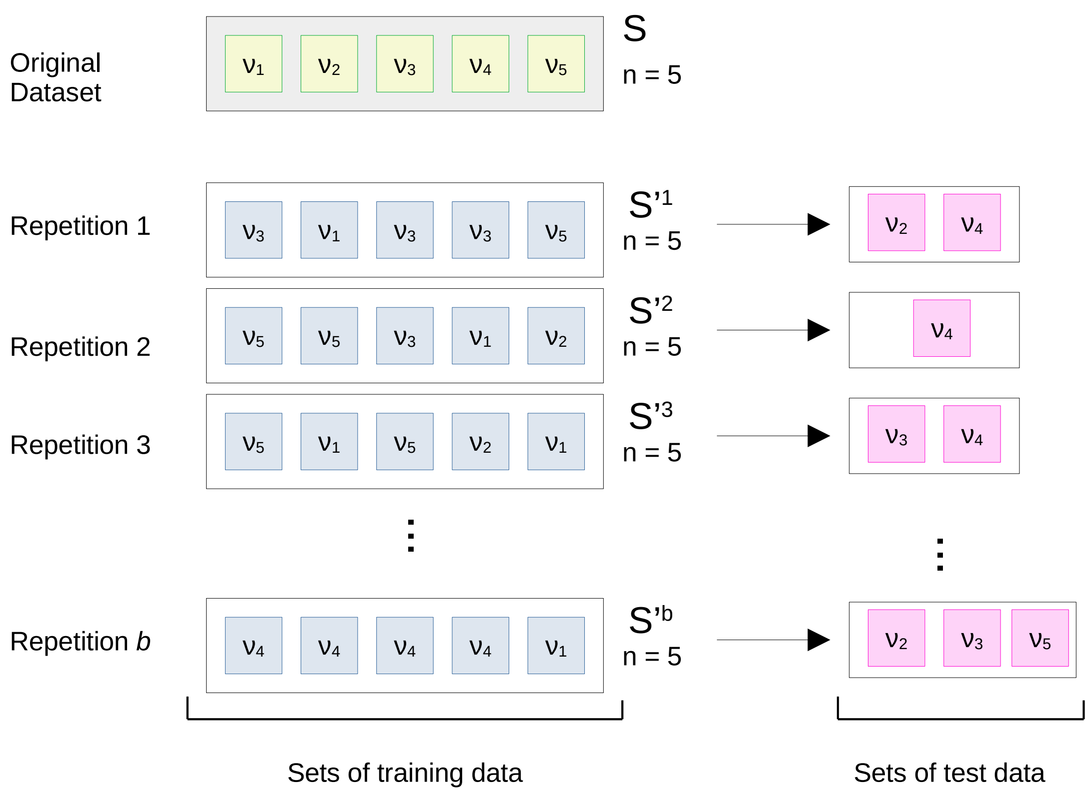
To illustrate the bootstrap approach, Figure 2.14 shows an initial sample S with sample size $n = 5$, where $S = \{\nu_1, \nu_2, \nu_3, \nu_4, \nu_5\}$. For the 1^{st} bootstrap repetition, the new sample S'^1 is $\{\nu_3, \nu_1, \nu_3, \nu_3, \nu_5\}$ which will be the training sample set. Consequently, $\{\nu_2, \nu_4\}$ is the test sample set as this is the remaining data. This procedure is repeated b times, producing b bootstrap datasets (HASTIE et al., 2009).

As this procedure carries out, an estimate α can be determined from it and for each bootstrap repetition, the estimation will be $\hat{\alpha}'^{b}$ and we can compute the bootstrap standard error $SE_b(\hat{\alpha})$ using Equation 2.5 as explained by James et al. (2021).

$$SE_b(\hat{\alpha}) = \sqrt{\frac{1}{b-1}\sum_{i=1}^{b}\left(\hat{\alpha}'^{i} - \frac{1}{B}\sum_{j=1}^{b}\hat{\alpha}'^{j}\right)^2} \tag{2.5}$$

This approach is used to estimate the accuracy of the Random Forest model (Section 2.4.4). The bootstrap has been used widely in case of unknown distributions (VRIGAZOVA, 2021).

Figure 2.14 - Bootstrap scheme.



A scheme of the bootstrap validation with the original sample S in green with sample size $n = 5$. From the original sample, for each bootstrap sample, there is a replacement of data so all of them have sample size $n = 5$ and are the training data. The test data is formed by the data not selected for the training.

SOURCE: Author.

2.5.1.5 Holdout - conventional approach

This is the conventional approach for remotely sensed data. In this scenario, the dataset is split into two subsets: one training and one test, which is not recommended for small datasets (THEODORIDIS; KOUTROUMBAS, 2009).

Theodoridis and Koutroumbas (2009) state that splitting the data into two subsets reduces the size of training and test data, which can be a drawback. Another point that they mention is what ratio to use for splitting. The ratio to split the data usually is 67% : 33% and 80% : 20% as they are analogous to bootstrap and 5-fold cross-validation, as explained by Hastie et al. (2009) and Lyons et al. (2018).

Additionally, the error inherent of this approach tends to decrease as the sample size n increases and for small test data, this approach may be unreliable (THEODORIDIS; KOUTROUMBAS, 2009).

2.5.1.6 Monte Carlo Simulation

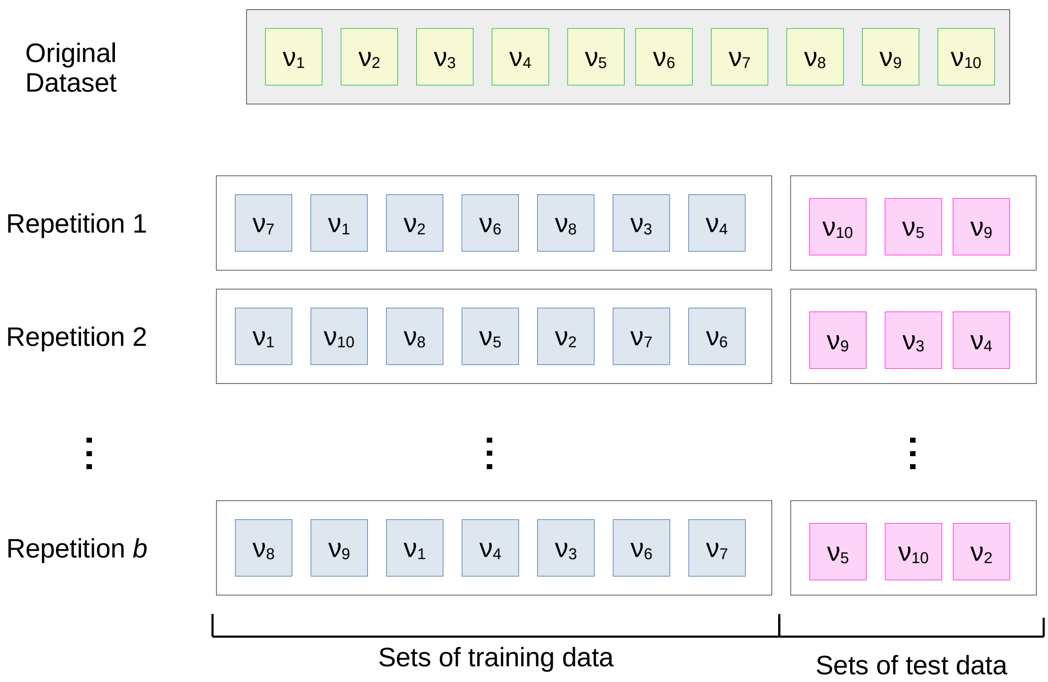
The Monte Carlo Simulation or Monte Carlo Cross-Validation was first proposed by Picard and Cook (1984). It is based on randomly splitting the data, as the holdout method, b times and "averaging the squared prediction errors over the splits" (SHAO, 1993). Additionally, the analyst defines the ratio to be used as well as the number of repetition (KUHN; JOHNSON, 2013).

Kuhn and Johnson (2013) point out that the number of repetitions affects the uncertainty of the performance estimates. In other words, the higher the number of repetitions b, the more stable is the estimate, being $b > 50$.

An example of the Monte Carlo Simulation scheme is in Figure 2.15, with $n = 10$ samples defined by the set $\{\nu_1, \nu_2, ..., \nu_{10}\}$. This set is randomly split into training and test sets in the ratio of 70% : 30%, respectively using simple random sampling.

Picard and Cook (1984) claim that this approach has a small computational cost with moderate and large datasets and with superior results when compared to other approach such as the bootstrap. Additionally, Haddad et al. (2013) concluded that this approach produced a more stable estimate compared to the leave-one-out approach.

Figure 2.15 - Monte Carlo Simulation scheme.



A scheme of the Monte Carlo Simulation using a dataset with sample size $n = 10$. From the original dataset, there is a simple random sampling to split the data into training and test data without any replacement in the ratio 70% : 30%, respectively.

SOURCE: Author.

2.6 Spatial Data Quality - measuring image classification accuracy

In order to minimise cartographic incoherences, it is important to identify them as well to assess as the data quality level (SANTOS, 2015). Taking this into account, the International Cartographic Association (ICA) in Guptill and Morrison (1995) proposed seven elements of spatial data quality where positional and thematic accuracies are two of them[4]. Additionally, the International Organisation for Standardisation (ISO) with the standard 19157:2013 defines six elements[5] for spatial data quality including thematic and positional accuracies, as well as completeness (ISO, 2013).

[4]The seven elements according to ICA are: lineage, positional accuracy, attribute accuracy, completeness, logical consistency, semantic accuracy and temporal information

[5]The six elements according to ISO 19157:2013 are: completeness, logical consistency, positional accuracy, thematic accuracy, temporal quality and usability element.

For assessing spatial data quality, there is the ISO 19157:2013 (ISO, 2013) as the international regulation and, in Brazil, the Technical Specification for Quality Control of Geospatial Data (ET-CQDG[6]) as an element of Brazil National Infrastructure for Spatial Data - INDE[7] (DSG, 2016).

Therefore, ISO 19157:2013 (ISO, 2013) and ET-CQDG (DSG, 2016) define Thematic Accuracy as the correct feature and attribute interpretation as well as its placement on the expected classes. Its subelements are: (i) classification correctness - comparison of the classes or features found on geospatial dataset with a model data (eg. ground truth); (ii) non-quantitative attribute correctness - evaluation of non-quantitative attributes when comparing them to attributes from the same features in a more accurate source and (iii) quantitative attribute accuracy - how close a quantitative attribute value can be to a value accepted or known to be true. Figure 2.16 illustrates an error of classification thematic accuracy.

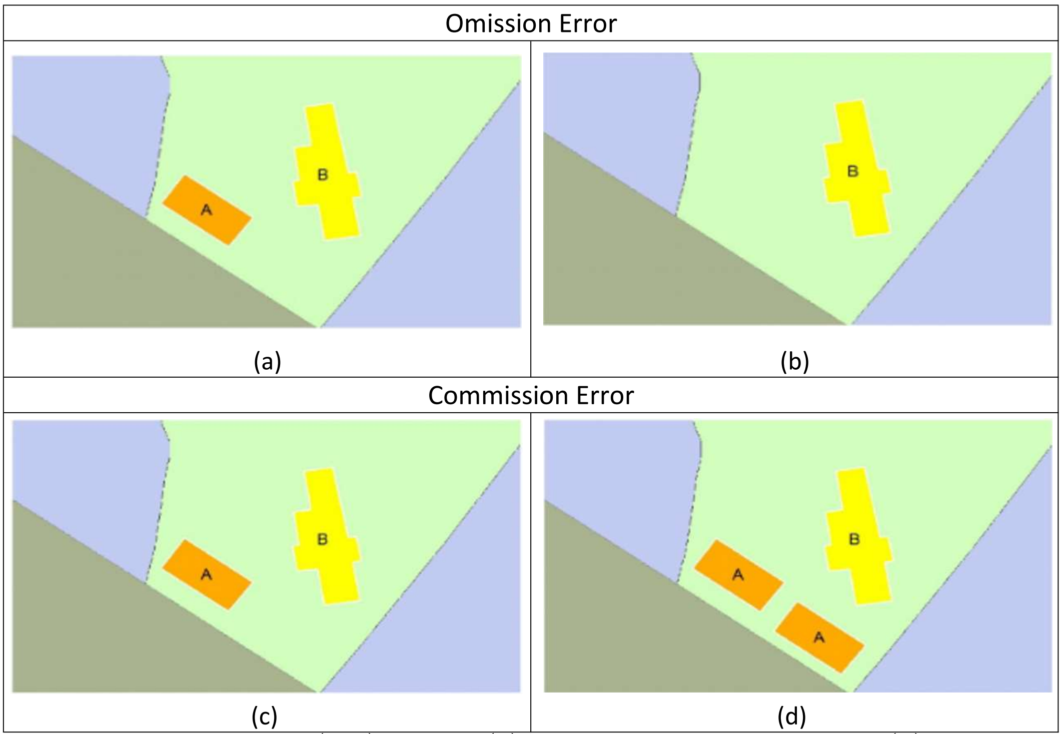
Figure 2.16 - Example of classification Thematic Accuracy.

Example of classification thematic accuracy where (a) is the ground truth and (b) is the classification result in the same area.

SOURCE: Adapted from DSG (2016).

[6]In Portuguese: Especificação Técnica para Controle de Qualidade de Dados Geoespaciais.
[7]In Portuguese: Infraestrutura Nacional de Dados Espaciais

Figure 2.17 - Example of Completeness - Omission and Commission errors.



Example of omission error (OE), where (a) is the reference data while (b) is the tested data with absence of information. Also, as examples of commission error (CE), (c) is the reference data and (d) is the tested data with more information than the reference data.

SOURCE: Adapted from DSG (2016).

Finally, Completeness is related to presence or absence of features in the geographic information (ISO, 2013; DSG, 2016). According to both standards, there are two subelements from completeness: commission and omission errors. Omission errors are related to the absence of features when compared to the reference data. Meanwhile, commission errors are related to the presence of double data or missclassified information. Both errors are shown in Figure 2.17.

2.6.1 Thematic accuracy

In order to understand how well the classification performed, ISO (2013) and DSG (2016) recommend the use of few metrics.

When it comes to Confusion Matrix, it was suggested by Congalton (1991), Foody (2002) and its structure is shown on Table 2.3. This matrix is formed by the same amount of rows and columns, expressing the quantity of certain category of a tested product (t, rows) with respect to a reference (r, columns), which this category can be either pixels or regions (DSG, 2016).

Table 2.3 - Confusion matrix structure.

Tested Unit	Reference Unit					Total
	1	2	...	r-1	r	
1	X_{11}	X_{12}	....	$X_{1,r-1}$	$X_{1,r}$	X_{1+}
2	X_{21}	X_{22}	....	$X_{2,r-1}$	$X_{2,r}$	X_{2+}
...	....	....	....	....	....	....
t-1	$X_{t-1,1}$	$X_{t-1,2}$	....	$X_{t-1,r-1}$	$X_{t-1,r}$	X_{t-1+}
t	$X_{t,1}$	$X_{t,2}$	....	$X_{t,r-1}$	$X_{t,r}$	X_{t+}
Total	X_{+1}	X_{+2}	....	X_{+r-1}	X_{+r}	N

SOURCE: Adapted from DSG (2016).

From this matrix, DSG (2016) mentions two metrics for classification accuracy, namely, overall accuracy (OA) and *kappa* index. The OA is expressed in percentage (Equation 2.6), based on the confusion matrix (Table 2.3).

$$OA = \frac{1}{N}\sum_{i=1}^{r} X_{i,i} \tag{2.6}$$

Where:
OA is the Overall Accuracy;
N is the total number of classified pixels;
r is the number of reference pixels;
$X_{i,i}$ is the number of corrected classified pixels from the i^{th} class.

Nonetheless, ISO (2013) recommend the use of *kappa* index, which is a discrete multivariate metric used for thematic accuracy assessment. It is also based on the confusion matrix (Table 2.3) and shown on Equation 2.7.

$$\kappa = \frac{N \cdot \sum_{i=1}^{t} X_{i,i} - \sum_{i=1}^{t} X_{i+} \cdot X_{+1}}{N^2 - \sum_{i=1}^{t} X_{i+} \cdot X_{+i}} \tag{2.7}$$

Where:
κ is the κ coefficient;

N is the total of tested pixels;
$X_{r,t}$ is the number of pixels in class r classified in class t;
X_{t+} is the amount of pixels classified in t class (test);
X_{+r} is the amount of pixels classified in the r class (reference).

Thematic Accuracy is understood as being part of the validation elements for classification; the other part is defined as Completeness and it is explained in the next Section (2.6.2).

2.6.2 Completeness

Taking into account completeness data, there are commission and omission errors (CE and OE, respectively). Commission Error relates to overcompleteness, when there is excess of data while Omission relates to incompleteness, when there is absence of data (VAN OORT, 2006). According to Congalton (1991), ISO (2013), DSG (2016). In general, their computations are given by Equations 2.8 and 2.9, respectively.

$$CE_i = \frac{\sum_{j=1}^{r} X_{ij}}{X_{i+}} \tag{2.8}$$

$$OE_i = \frac{\sum_{j=1}^{t} X_{ji}}{X_{+i}} \tag{2.9}$$

Where:
r is the reference unit;
t is the testes (or predicted) unit;
CE_i is the commission error for class i;
OE_i is the omission error for class i;

Other used metrics for remotely sensed data are user accuracy (UA) and producer accuracy (PA). They are presented in Equation 2.10 and Equation 2.11 and are the complement of CE and OE, respectively. Studying accuracy instead of error may bring some easiness for data interpretation, depending on the used situation.

$$UA_i = 1 - CE_i \therefore UA_i = \frac{X_{ii}}{X_{i+}} \tag{2.10}$$

$$PA_i = 1 - CO_i \therefore PA_i = \frac{X_{ii}}{X_{+i}} \tag{2.11}$$

All are usually expressed in percentage. These errors aid the Thematic Accuracy for analysing the classification accuracy and are part of the validation process in Remote Sensing data.

3 PROPOSED APPROACH

In this chapter we will formally propose an approach for computing and defining the quality of reference data for supervised image classification aiming to answer the question: *what is the influence of the reference data quality for classification modelling, results and accuracy?*

The quality of a pixel can be related to some questions, such as: does the pixel present only the spectral response of a single class? Is it mixed to other classes? If so, how much is it mixed? To answer these questions, in this study we use a higher resolution classified image, entitled here as baseline image classified image, or HR, applying the following conditions:

I. The baseline classified image is co-registered to the lower resolution image (LR). In the case the images are not properly co-registered, there may be misplacement of information leading to a gross errors during the validation process;

II. The set of classes defined for HR and LR are the same. This condition aids interpretation on this initial study although changes in this condition may be applied in further studies;

III. The HR spatial resolution should be at least six times better than LR resolution in a way that there must be at least 36 HR pixels inside a LR pixel so the minimum number of samples for defining estimates with certain stability and confidence;

IV. The baseline image thematic accuracy is considered very high, with OA $\geq$ 85% or *kappa* $\geq$ 0.80. A reason for that is that errors in the baseline classification may affect directly the results therefore errors in this part should be avoided;

V. The reference sample will be selected if the modal class proportion of HR pixels within the LR pixel is equals to or greater than 50%, i.e., $prop \geq$ 50%.

To illustrate the scenario with the conditions, Figure 3.1 shows an image from the MultiSpectral Imager (MSI) sensor onboard Sentinel-2 satellite in two formats: original $10m$ spatial resolution in true colour composition as a higher resolution image (HR) and the same image resampled (using average) to $60m$ as the lower resolution image (LR). These images, Figure 3.1(a) and Figure 3.1(b), show a visual example of the difference in spatial resolution. These first two images indicate that a lower resolution image is likely to have mixed pixels regarding classes such as Bare Soil, Forest and Water. Subsequently, Figure 3.1(c) presents the baseline classified image with the classes Bare Soil (squares), Forest (triangles) and Water (circles). Figure 3.1(d) presents the selected reference data, which is the Water class.

Figure 3.1 meets all five conditions. As both images are the same, though one is re-sampled, they are perfectly registered, meeting condition I. Both images are studied using the same classes (condition II). Third, this is simulated scenario, hence the thematic accuracy of the baseline image, Figure 3.1(c), is perfect, meeting condition III. Also, there are 36 HR pixels inside a LR pixel, meeting condition IV.

When it comes to pixel quality, in condition V, for the case of Figure 3.1 the reference data is selected if, and only if, it has at least 90% of HR pixels within it from the same class, i.e. the modal class proportion is $\geq$ 90%. Therefore, from the four LR pixels, only the bottom-right pixel is selected as reference sample with modal class proportion $prop = 91.67\%$. The remaining pixels present 83.3% (upper-left), 55.5% (upper-right) and 69.4% (bottom-left). Thus, even though not all pixels are selected as reference data, all of them match condition V.

Once the pixel quality is defined, we can count the number of samples per class and according to their quality, so a histogram can be created. This histogram can aid us to understand how each class is distributed, computing measures of central tendency. In other words, from the histogram, we can study about the class representativity. As we have grouped data, for each class, the mean grouped value ($\overline{g}$) and its respective sample variance (s^2) are defined by Equation 3.1 and Equation 3.2. From the variance, the standard deviation (std) is computed as being the square root of the variance ($std = \sqrt{s^2}$).

$$\overline{g} = \frac{\sum_{i=1}^{M} g_i f_i}{\sum_{i=1}^{M} f_i} \tag{3.1}$$

$$s^2 = \frac{1}{\sum_{i=1}^{M} f_i - 1} \left[\sum_{i=1}^{M} g_i^2 f_i - \frac{\left(\sum_{i=1}^{M} g_i f_i \right)^2}{\sum_{i=1}^{M} f_i} \right] \tag{3.2}$$

Figure 3.1 - Manual extraction of lower resolution reference pixels based on higher resolution pixels.

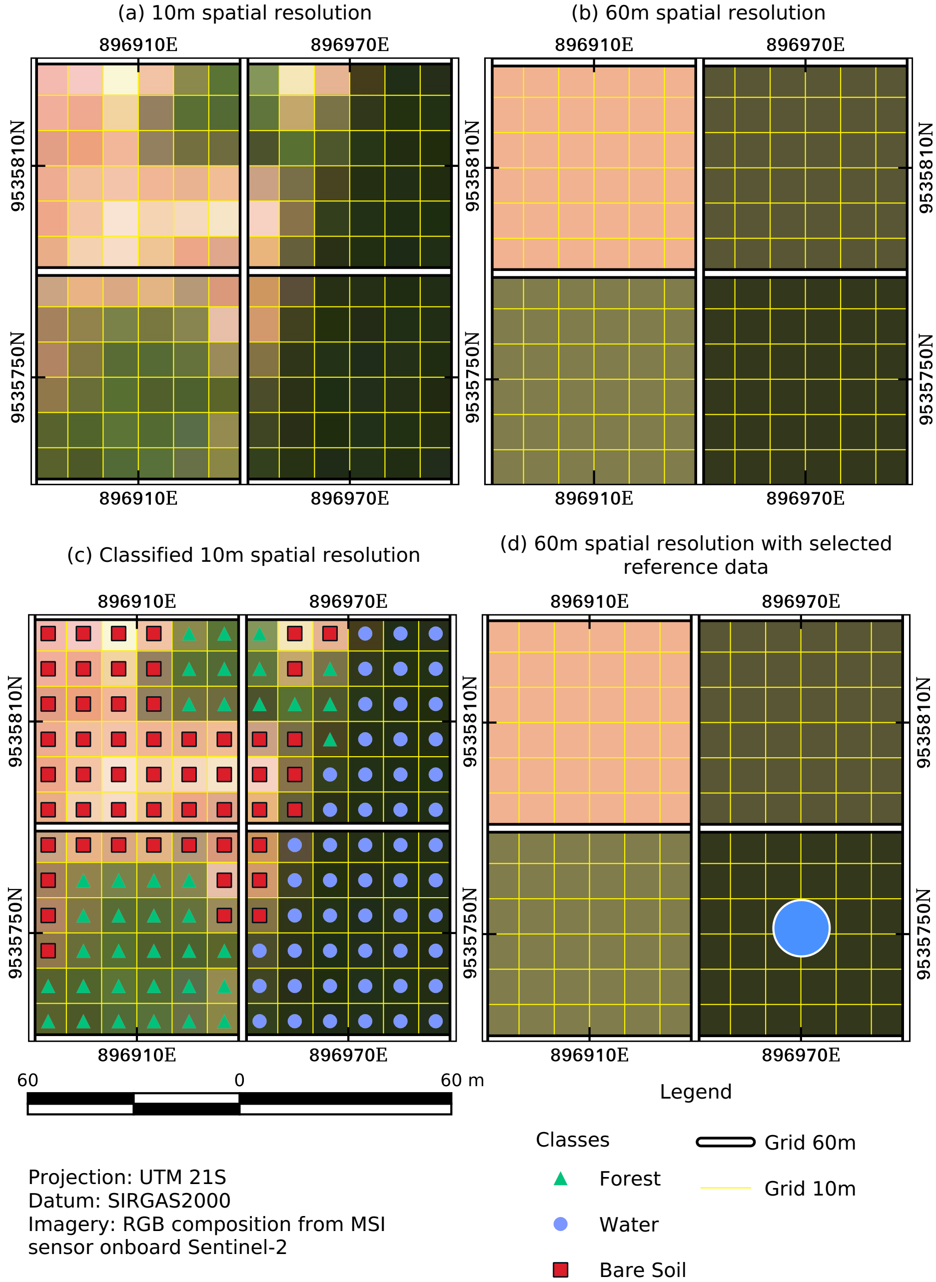


Illustration of reference data selection using an original MSI Sentinel-2 scene with (a) 10m spatial resolution as higher resolution image; (b) resampled to 60m spatial resolution as lower resolution image (LR); (c) baseline classified image; (d) LR image with selected reference data assuming the selected majority of elements ($\geq 90\%$).

SOURCE: Author.

Where,

i is the group index, varying from 1 to the total number of groups M;

g is the group centre;
f is the frequency of each group.

We can also study about the probability distribution of the histogram using skewness and kurtosis of the grouped data. According to Shanmugam and Chattamvelli (), skewness "summarises the degree of asymmetry of a unimodal distribution", when this value is positive, the mass is concentrated on the left of the chart; when the value is negative, the mass concentration is of the right part of the chat and if it is zero, then the curve is symmetric. On the other hand, the authors claim that kurtosis "characterises the accumulation of probability mass toward the centre"; the distribution can be (i) mesokurtic, where the is no excess[1] and it is related to the normal distribution; (ii) leptokurtic, where there is positive excess or (iii) platykurtic, where there is negative excess. The equations to define these two measures are Equation 3.3 and Equation 3.4.

$$skewness = \frac{\sum_{i=1}^{M} (g - \overline{g})^3}{\sum_{i=1}^{M} f_i \cdot \left(\frac{\sum_{i=1}^{M} (g - \overline{g})^2}{\sum_{i=1}^{M} f_i} \right)^3} \tag{3.3}$$

$$kurtosis = \frac{\sum_{i=1}^{M} (g - \overline{g})^4}{\sum_{i=1}^{M} f_i \cdot \left(\frac{\sum_{i=1}^{M} (g - \overline{g})^2}{\sum_{i=1}^{M} f_i} \right)^4} \tag{3.4}$$

After verifying the histogram, we verify the impact of its variation when measuring the thematic and classification accuracies. For doing so, the first step is splitting the reference data into training and test samples. The training reference data is divided into two categories: (i) range quality, when groups of samples are clustered in 5% range, indicated here as $[prop, prop{+}5\%[$ and (ii) accumulated quality, when samples are clustered into a defined minimum multiples of 5% until 100% quality, indicated as $[prop, 100\%]$[2]. On the other hand, the test samples are defined in three sets: (i) the first summarises the quality of an entire image using modal class proportion in the set $[50\%, 100\%]$; (ii) the second represents only pure pixels in the set $[100\%]$

[1]Excess kurtosis is defined by kurtosis - 3. For more information, please refer to Shanmugam and Chattamvelli ().

and (iii) the last tests the classifier *per se* varying the quality of pixels in the set $[prop, 100\%]$. A summary of the sets is presented in Table 3.1, and combining all options result in six different setups of experiment.

Table 3.1 - Sets of training and test data.

Train Samples	Test Samples
$[prop, 100\%]$	$[50\%, 100\%]$
$[prop, prop + 5\%[$	$[100\%]$
	$[prop, 100\%]$

SOURCE: Author.

These six setups are justified and explained in Table 3.2, showing the combinations and the objectives of each of them, generalising situations where the aim is to test either the classifier or the map. In this study, testing the image/map concerns using the representation of all possible pixels of classes, which can have mixed pixels patterns. Contrarily, when testing the classifier, it is expected to input purer pixels. Therefore, as there are two different objectives, the training and test sets combinations aim to explore their outcomes for both situations in test and training. In order to study these elements, three perspectives are analysed:

(i) the response of each used classifier to the variation of quality of the reference data;

(ii) how the convergence of reference data quality affects the classification assessment and

(iii) the effect of quality training and test samples separately on the classification assessment.

Another point of view of it is the study of the analyst evolution and its effects on the image classification assessment. For this, Setup 3, 4 and 6 are used, being the latter the most commonly used. Setup 6 shows the evolution of a least experienced analyst up to a more experienced one for collecting reference data and for this reason is presented in Figure 3.2. All other Setups feature spaces are shown in Appendix A.

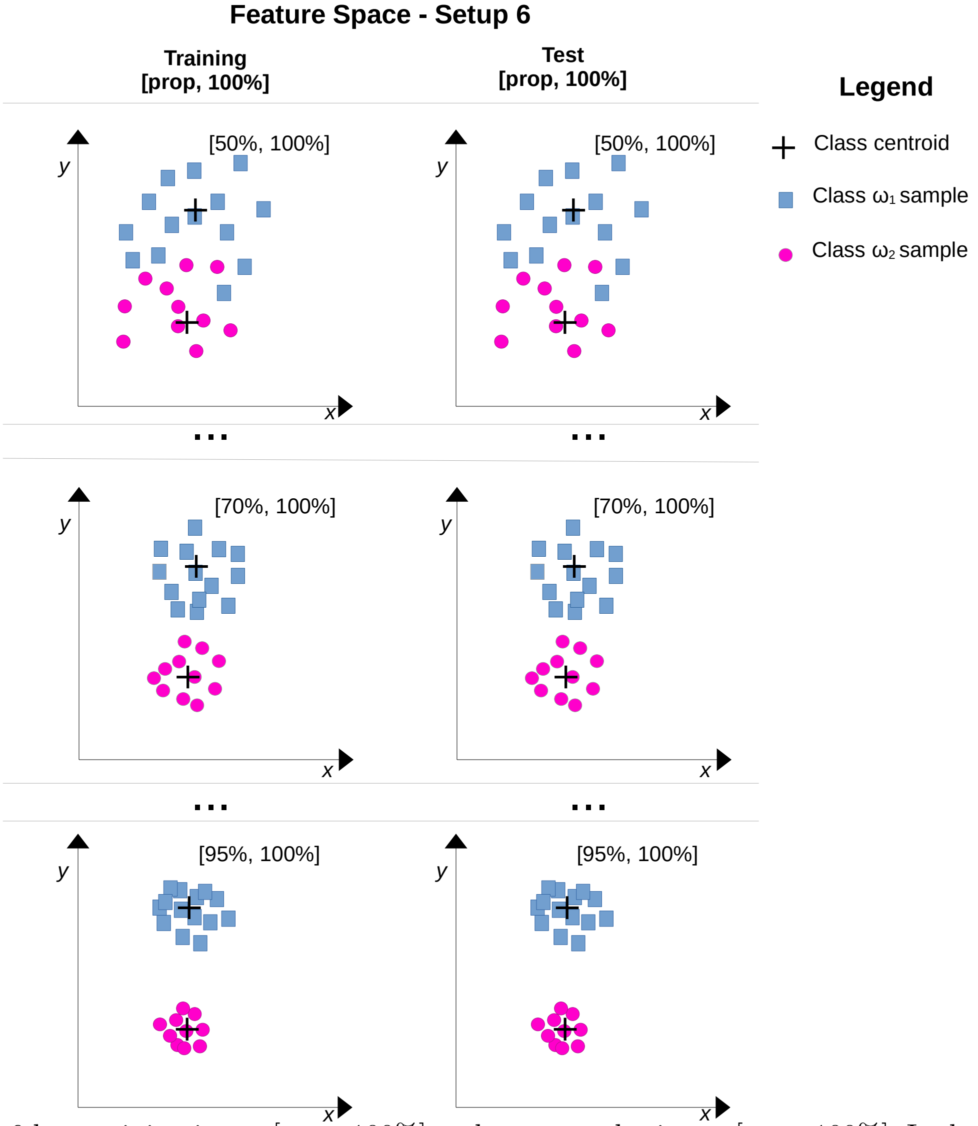
[2]The notion $[prop, prop + 5\%[$ indicates closed interval at *prop*, i.e. it considers the value *prop*; open interval at $prop + 5\%$ indicating that this exact value is not used in the set.

Table 3.2 - Setups combinations when splitting training and test samples.

Setup	**Train Set**	**Test Set**	**Simulation**
1	$[prop, prop + 5\%[$	$[50\%, 100\%]$	Tests the performance of increasingly purer sets of training samples though in a limited range against a generalised test set, which is supposedly representative of typical modal class areas.
2	$[prop, prop + 5\%[$	$[100\%]$	Tests the performance of increasingly purer sets of training samples though in a limited range against a pure set of pixels within the modal class. Idealised test sets are the ones supposedly chosen in supervised tasks.
3	$[prop, 100\%]$	$[50\%, 100\%]$	Tests the performance of classifiers estimated with increasingly purer training samples against a generalised test set, which is supposedly representative of typical modal class areas.
4	$[prop, 100\%]$	$[100\%]$	Tests the performance of classifiers estimated with increasingly purer and complete pure sets of training samples against an idealised test set.
5	$[prop, prop + 5\%[$	$[prop, 100\%]$	Tests the performance of increasingly purer sets of training samples in a limited range against a generalised test set, which is supposedly representative of typical modal class areas.
6	$[prop, 100\%]$	$[prop, 100\%]$	This is the scenario of common reference data acquisition, where the analyst acquire both training and test data according to their collecting experience.

SOURCE: Author.

Figure 3.2 - Illustration of Feature Space for Setup 6.



Setup 6 has training in set $[prop, 100\%]$ and test samples in set $[prop, 100\%]$. It shows the feature space varying according to the increase of pixel quality.

SOURCE: Author.

4 EXPERIMENTAL PLANNING

Once the theoretical idea is presented, in order to proceed with the development of the methodology, the experiment is done in a controlled situation to study the effect of samples quality when the image to be classified is perfectly registered to the baseline image which is to be obtained by a simulation.

In a nutshell, the algorithm uses the higher-resolution image (HR) together with the lower-resolution image(LR) and, for each LR pixel, it evaluates the number of HR cells inside a LR cell and determines whereas that LR cell will be selected or not as reference data. Figure 3.1 illustrates this approach.

Hence, by analysing a classified HR, the algorithm can determine which LR cells have an acceptable proportion of the modal class. A LR cell is selected if and only if the defined modal class proportion ($prop$) of HR cells from the modal class within it, in equals to or greater than 50%. Once the algorithm is finished selecting candidate samples, it can proceed to the LR classification and validating it (i.e. assessing thematic accuracy and completeness), presented in Section 4.2.

In order to move on to to studying the quality of reference data, the used materials for the stages are presented in Section 4.1.

4.1 Materials

The materials used for this study were softwares and programming environments as well as satellite imagery. An important factor considered was the open-source policy so this study can be easily reproduced.

Another point is that the author considered a colourblind friendly policy, therefore the majority of images are presented in a manner that colourblind people can easily read; if not, images in black & white (BW) format are presented in the Appendix C.

4.1.1 Study area

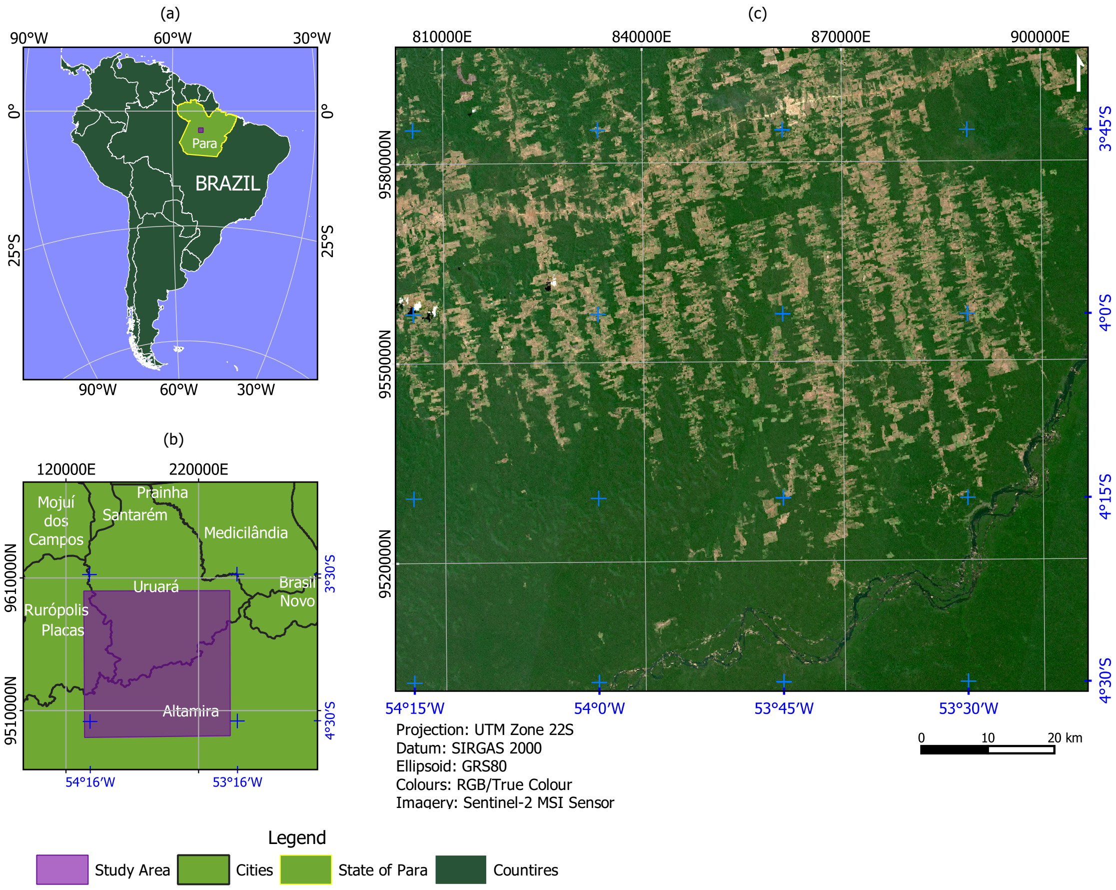
The study area lays in the Brazilian Amazon, located in Pará State, in the mesoregion of Sudoeste Paraense and microregion of Altamira. Its extents are: 3º36'48"S to 4º36'12"S, and 53°18'40"W to 54º21'0"W in geographical coordinates, Datum SIRGAS 2000. The area is approximately $12,068km^2$. Most of the study area belongs to the municipality of Altamira while the remaining area covers Placas and Uruará municipalities, with their respective urban areas, as shown in Figure 4.1.

Besides, Alvares et al. (2013) state that, according to Köppen classification, the study area has a Tropical climate, mainly the AF type, and it presents an annual rainfall between $2,200$ and $2,700mm$, and a mean annual temperature greater the 26°C. Regarding its hydrography, the area is located within the Amazon River basin, in the Xingu Paru sub-basin.

The diversity of geographical characteristics made this area suitable to apply the methodology proposed. The presence of urban areas along the *Transamazônica* Highway is contrasted with *Iriri* river floodplain and a part of two protected areas: Extractive Reserve of Iriri river and Ecological Station of *Terra do Meio*. Furthermore, as Pará is one of the Brazilian States that have presented greater deforestation area in 2019 (ASSIS et al., 2019; INPE, 2019), a large part of the study area presents fishbone deforestation pattern.

With the study area set, the imagery and it information are presented afterwards.

Figure 4.1 - Study area - Amazon Rainforest.



(a) South America and its countries boundaries, showing the State of Para in a lighter colour and the study area. (b) The State of Para with the cities where the Study Area lays in. (c) Study area in True Colour composition from Sentinel-2 MSI sensor taken in 09 August 2020.

SOURCE: Adapted from ESA (2015a) and IBGE (2020).

4.1.2 Used land cover classes for image classification

Nearby areas were studied by Reis (2014), Reis et al. (2018), Reis et al. (2020), Reis et al. (2020) and Soares et al. (2020) and they all reported Land Cover trajectories over time. Their studies used *in situ* data and remote sensing imagery for defining the labels. Another study is Coutinho et al. (2013), which was used as base for the other mentioned studies that used classes that will be summarised in this study according to Table 4.1. The classes were summarised because when different classes present same elements, they tend to present high confusion rate (REIS et al., 2018).

Table 4.1 - Land cover classes used for this research.

Class	Description
Bare Soil	Exposed soils, with close to zero vegetation.
Crops	Cultivation, Litter or Shrubs
Forest	Mature Forest, or areas with dense forests and trees.
Water	Water mass

SOURCE: Reis et al. (2018), Reis et al. (2020).

4.1.3 MultiSpectral Imager (MSI) sensor on-board Sentinel-2

The MultiSpectral Imeger (MSI) sensor on-board Sentinel-2-A was launched in 2015 by the European Spatial Agency (ESA) and its twin satellite, Sentinel-2-B, was launched in 2017, hence being a mission composed by these two satellites (ESA, 2015b). The MSI sensor has twelve bands, as can be seen in Table 4.2. ESA (2015b) points out some features from the MSI sensor, for example that it has bands on the visible (VIS), infrared region as well as near infrared (NIR), vegetation red-edge and short-wavelength infrared (SWIR) bands.

Such bands, together with a spatial resolution of $10m$ (VIS, NIR), $20m$ (SWIR and vegetation red-edge) and $60m$ (SWIR and water vapour), as shown in Table 4.2, can allow the monitoring of the land, oceans, emergency management and security, as stated in ESA (2015b).

The Sentinel-2 mission covers continental land surfaces between latitudes $56°S$ and $84°N$. When it comes to revisit time (temporal resolution), as the mission consists of twin satellites, the revisit time is 5 days -10 days for each satellite. It has a sun-synchronous orbit, being the satellites with 180° from each other and both with inclination of 98.62° (ESA, 2015b).

Their product levels are: Level-0, Level-1A, Level-1B, Level-1C and Level-2A, being the latter two available for users. According to ESA (2015b), Level-1C consists of orthorectified reflectance from the Top-of-Atmosphere (TOA) while Level-2A provides reflectance from the Bottom-Of-Atmosphere (BOA).

Table 4.2 - Sentinel-2 MSI sensor bands.

Band	**S2A**		**S2B**		**Spatial Resolution** (m)
	Central Wavelength (nm)	**Bandwidth** (nm)	**Central Wavelength** (nm)	**Bandwidth** (nm)	
1 - Coastal aerosol	442.7	21	442.3	21	60
2 - Blue	492.4	66	492.1	66	10
3 - Green	559.8	36	559	36	10
4 - Red	664.6	31	665	31	10
5 - Vegetation red-edge	704.1	15	703.8	16	20
6 - Vegetation red-edge	740.1	15	739.1	15	20
7 - Vegetation red-edge	782.8	20	779.7	20	20
8 - NIR	832.8	106	833	106	10
8A - Vegetation red-edge	864.7	21	864	22	20
9 - Water vapour	945.1	20	943.2	21	60
10 - SWIR Cirrus	1373.5	31	1376.9	30	60
11 - SWIR	1613.7	91	1610.4	94	20
12 - SWIR	2202.4	175	2185.7	185	20

SOURCE: Adapted from ESA (2015b).

For this study, the sensing date was 09 August 2020, at UTC 14:01:01. Its processing baseline number (N) is 02.14, its relative orbit number (R) is 067 and it tile number field (T) is 21MZR. The processing level is 2-A, which corresponds to Bottom-Of-Atmosphere (BOA) corrections. Cloud percentage of this image is 1.18% (0.018318). The image in its true colour composition (TCI) is shown in Figure 4.1(c).

The number of pixels in the image with $10m$ spatial resolution was 10890 x 10890. The bands used were Blue (B02), Green (B03), Red (B04) and Near-Infrared (NIR, B08).

As this is a controlled situation, the original image was considered the higher-resolution image and this image was resampled using mean values to 60m spatial resolution to be the lower-resolution image (LR). Hence:

- Higher Resolution: Sentinel-2 MSI 10m (original);
- Lower Resolution (LR): Sentinel-2 MSI resampled to 60m, using average.

In case of clouds, as there was not a substantial percentage of them, these pixels were excluded manually so they would not interfere in any further computations.

4.1.4 Software and used data

All the coding process was done in Python 3.7 environment (ROSSUM; DRAKE, 2009) with the packages shown in Table 4.3. Aside from the coding processes, for Geographical Information Systems (GIS) manipulations, Quantum GIS (QGIS) 3.16 (QGIS DEVELOPMENT TEAM, 2009) was used. Moreover, for object-based segmentation with Multiresolution Segmentation, *eCognitio*n 9.1 (Trimble Germany GmbH, 2014) was used and region-based classification was done in TerraView 5.5.1 (INPE, 2020).

All images used in this study were free of charge and can be found at the Sentinel Hub Copercinus[2] website, from the ESA. The used sensors are described in Section 4.1.3.

[1]Accessible at: `https://scihub.copernicus.eu/dhus/#/home`

Table 4.3 - Python 3.7 packages used.

Name	Description	Version
NumPy	Scientific computing	1.19.2
Rasterio	Access and manipulate raster data	1.1.0
Pandas	Data structure and analysis tools	1.2.3
Geopandas	Geospatial vectorial data manipulation	0.8.1
Sci-kit learn	Predictive data analysis	0.24.1
SciPy	Scientific computing	1.1.6
Matplotlib	Create visualizations	3.3.4
Plotly	Create graphs	4.14.3

SOURCE: Hunter (2007), Pedregosa et al. (2011), Gillies et al. (2013–), Jordahl (2014), PLOTY TECHNOLOGIES INC. (2015), Harris et al. (2020), THE PANDAS DEVELOPMENT TEAM (2020), Bell et al. (2021).

4.2 Reference Sample Selection - RSS

The proposed semi-automated method for defining the quality of reference samples in a remote sensing image and assessing their effect on image classification is entitled Reference Sample Selection (RSS). In this section, we explain step-by-step of how this procedure is semi-automatically conducted in a controlled situation. Figure 4.2 presents a broader visualisation of the process according to how they are presented in this manuscript. In a general form, the process is defined as follows:

i **Baseline Classification**: classification of the higher-resolution image presented in Section 4.2.1. This step firstly classifies the HR image using two approaches: (i) supervised pixel-by-pixel Random Forest classifier and (ii) region-based segmentation followed by supervised Decision Trees C5.0 classification. This step generates the baseline image (HR) that is later used for RSS Part I. Also, at this point, Conditions I, II, III and IV[3] must have been met;

ii **Reference Sample Selection (RSS) part I**: which is divided into two parts, both presented in Section 4.2.2:

a) Acquisition of modal class proportion (*prop*) of each cell of the corregistered lower-resolution image grid;

b) Selection of samples in LR image where the modal class proportion meets Condition V, i.e. $prop \geq 50\%$;

c) Analysis of histogram of the distributed data.

iii **Reference Sample Selection (RSS) Part II**: which is also divided into two parts and more detailed explained in Section 4.2.3:

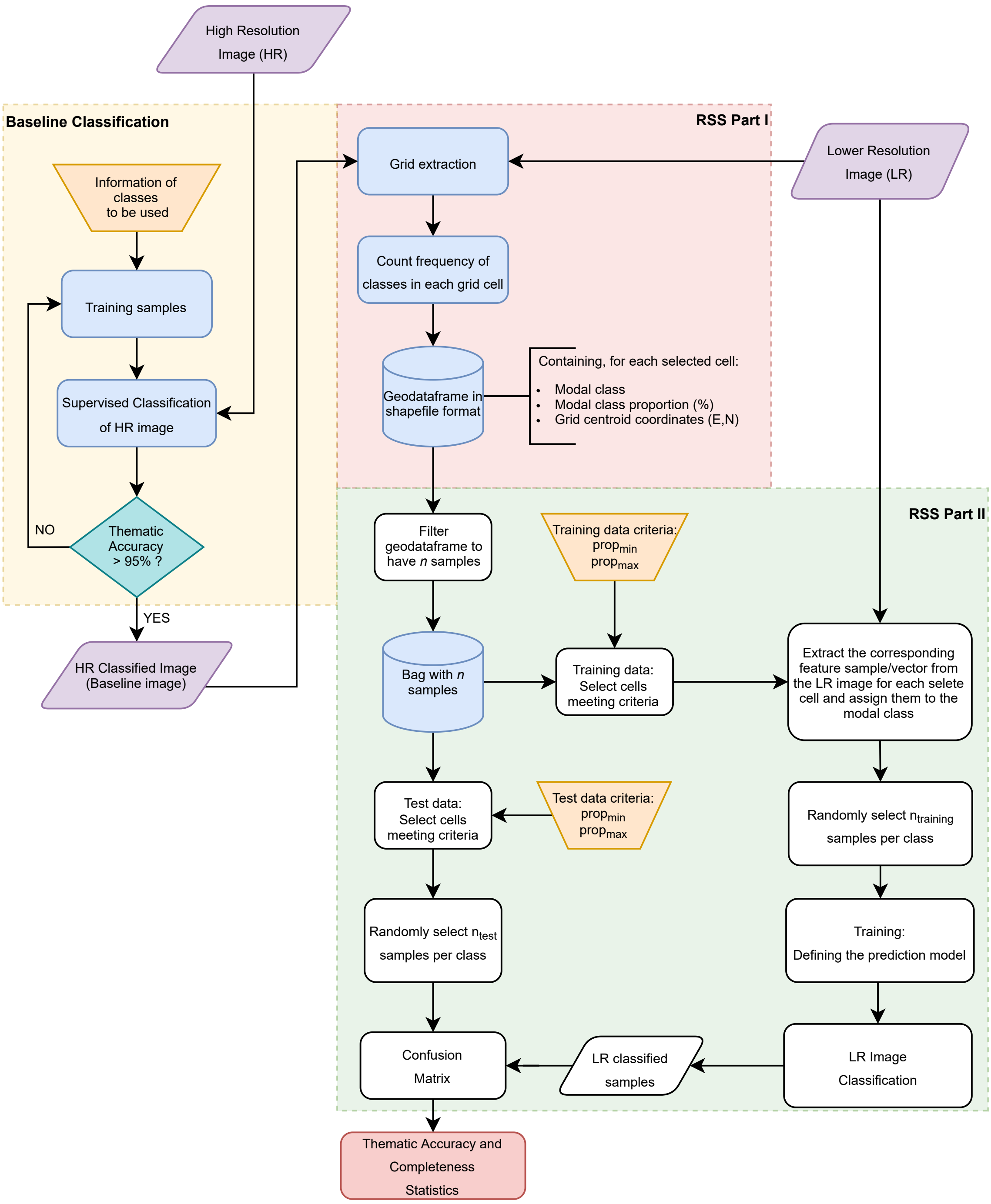
a) Usage of subset of acquired samples to train the classifiers;

b) Supervised classification of the LR image using Monte Carlo Simulation;

iv **Spatial Data Quality**: Accuracy computation and analysis, exposed in Section 4.2.4

[3]All Conditions are presented in Section 3, page 39.

Figure 4.2 - Flowchart for Reference Sample Selection (RSS) process.



This flowchart is divided into Baseline Classification and RSS, as the subsequent subtopics of this chapter. White boxes refer to processes to be repeated in the Monte Carlo Simulation

SOURCE: Author.

4.2.1 Baseline classification

The baseline Classification is the supervised image classification of the higher-resolution image. This step directly affects the quality of the final classifications, therefore it must be done thoroughly. We pointed out that the classes in the baseline classification and for the RSS process are the same thus meeting Condition II.

The higher-resolution image is classified using two approaches: (i) pixel-based supervised Random Forest classification, (ii) region-based supervised Decision Trees C5.0 classification. There are different classifiers due to the availability of using a classifier with segmentation in a software and in a Python environment.

For the Pixel-based classification, the Random Forest classifier is set using Python 3.7 environment (ROSSUM; DRAKE, 2009) with the packages Numpy (HARRIS et al., 2020) and Scikit-Learn (PEDREGOSA et al., 2011). For this classifier, the number of trees was set to 100, the Out-Of-Bag (OOB) score was used with a random state of 9999. It was also set to a minimum number of 2 splits and 1 leaf. Due to the image size, the HR image was cropped into four parts. The prediction model was defined using the entire image and the classification was applied separately to each cropped part. In this baseline classification, the same type of object is used for training and test data: pixels. This way, the reference data is split into 70%/30% for training and test, respectively.

Meanwhile, for the segmentation, the used metrics were Mean, Standard Deviation, Geometric ones, such as Area, Length, Border Length, Number of Pixels, Relation Border to Image border, Thickness, Volume, Width, Asymmetry, Density and Texture metrics, like Homogeneity, Contrast, Dissimilarity and Entropy. Once the image was segmented, the region samples for training were selected manually. Regarding test samples, they were randomly selected as 30% of total Pixel-based samples which leads to a greater number of test samples than training samples.

In both cases, the Thematic Accuracy and Completeness - the validation process - must be $OA \geq 85\%$ (Condition IV); if possible, the analyst should correct the misplaced classified pixels afterwards. An explanation for that is that the LR classification quality is affected by the Baseline Classification and errors inherent of this first process are likely to bias any further decision-making process.

Once the Baseline Classification is conducted, the baseline image is generated and it is possible to carry out to the extraction of reference data of LR based on the baseline image with the RSS process.

4.2.2 Reference Sample Selection part I - filtering candidate samples

This part, as presented in Figure 4.2 there is the grid extraction of both HR and LR images. Then the algorithm counts the number of classified baseline cells within the LR cell, defining the modal class and its frequency (modal class proportion).

This process follows the geographic extents of the LR cells, forming a mask with the size of a LR cell, regarding its spatial resolution in x and y axes. The mask is used to delineate the crop extents of the baseline image for each analysed LR cell, as shown in Figure 4.3.

This grid mask runs along all LR cells within the cropped area and, for each LR cell, the modal class, its respective proportion and the LR cell centroid coordinates (Easting and Northings) are computed and stored in the geodataframe. As this is a vectorial grid mask, it considers explicitly the HR cells inside of it, regarding intersected cells to the mask. Considering the possibility to apply this algorithm to a real scenario, the vectorial characteristics of the grid mask overcome possible registration errors or non-integer number of HR within a LR cell, both likely situations. Hence, even though this algorithm consumes a greater processing time and has greater computational effort, it was chosen for this study.

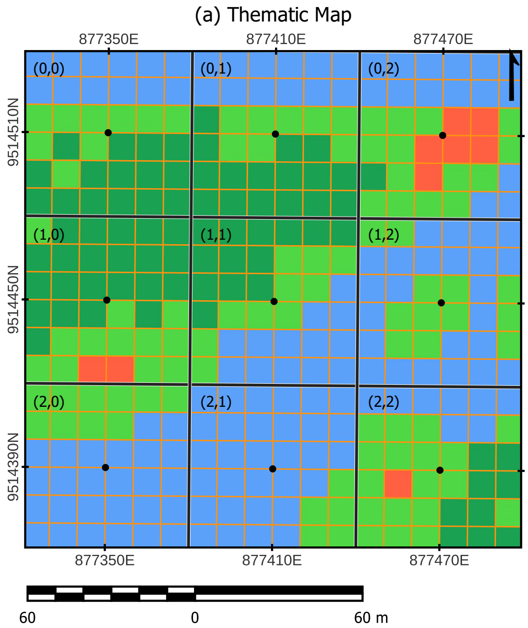
To illustrate how the grid mask runs through the baseline image, Figure 4.3 shows the first four cells runs to acquire the LR cell information so a geodataframe is created. Once the mask runs the entire baseline image acquiring information, the geodataframe is exported in a in a *shapefile*[5] format.

The geodataframe in a shapefile format was chosen so it could be used for GIS analysts in common GIS softwares and it also has vectorial information over the database. This geofdataframe is later transformed into raster information if GeoTIFF[6] format with bands informing the modal class and modal class proportion. Therefore, the geodataframe is formed by candidate cells to be reference data for image classification, as all of them have $prop \geq 50\%$, meeting Condition V.

[5]ESRI Shapefile format is a vectorial data storage format. It stores the position, shape and attributes of geographic features. Please refer to ESRI (2020) for more information.

[6]GeoTIFF is a format extension that stores geocoding and georeference informations in a raster file by tying a raster image to a known map projection or model space. For more information, please refer to Ritter and Rith (2000).

Figure 4.3 - RSS part I - running the grid mask through the LR image.



(b) Geodataframe

Pixel	Modal Class	prop	Centroid (E,N)	Candidade Sample
(0,0)	Forest	41.67%	(877350, 9514510)	No
(0,1)	Forest	44.44%	(877410, 9514510)	No
(0,2)	Water	41.47%	(877470, 9514510)	No
(1,0)	Forest	61.11%	(877350, 9514450)	Yes
(1,1)	Water*	36.11%	(877410, 9514450)	No
(1,2)	Water	63.89%	(877470, 9514450)	Yes
(2,0)	Water	72.22%	(877350, 9514390)	Yes
(2,1)	Water	86.11%	(877410, 9514390)	Yes

*In the case of more than one modal class, the algorithm choses the smallest number. In order, Water refers to 1 and Forest refers to 2. So Water is selected.

Legend

Classes

- Water
- Forest
- Crops
- Bare Soil

Features

- Grid 60m
- Grid 10m
- Pixel centorid

Cartographic Information

Datum: SIRGAS2000
Ellipsoid: GRS80
Coordinate Reference System: UTM 22S
Scale: Indicated

The grid mask runs through all LR pixels in order to acquire the modal class and the modal class proportion and add them to a geodataframe. (a) is the Thematic map or baseline image containing pixel row/column number and (b) is the attribute table of the geodataframe indicating which pixels are selected as candidate samples. This image in Black and White is Figure C.1, referred to in Appendix C, page 197.

SOURCE: Author.

From the generated geodataframe, it is possible to analyse the distribution of candidate data (*prop* $\geq 50\%$) among the used classes, so it is possible to understand the class representativity regarding sample quality.

At this point, all Conditions should be met to move on to RSS Part II, which is the process of selecting the training and test data as well as the image classification using Monte Carlo Simulation.

4.2.3 Reference Sample Selection part II - selecting reference data and image classification

In this part of the processes, as presented in Figure 4.2, the candidate samples are filtered to a *bag*. Then, the samples into the *bag* are selected as reference samples so it is possible to carry on to the image classification using the Monte Carlo Simulation.

In this sense, a stratified random sampling is used.

The first part is to filter the data to the *bag*, a process occurred once per Setup. Its main goal is to set the same number of samples per stratum, $n_{samples}$, so the probability of randomly selecting samples from all strata is the same. The strata are divided per class and per intervals of $[prop, prop + 5\%[\ \forall\ prop \in \{50\%, 55\%, 60\%, 65\%, 70\%, 75\%, 80\%, 85\%, 90\%, 95\%, 100\%\}$, as shown in Figure 4.4.

The number of samples per stratum ($n_{samples}$)is defined according to the less occurred stratum and can be rounded as it best fit the analyst. This way, the total amount of sample in the *bag*, n_{bag}, is defined by the number of strata per class, 11 (Figure 4.4), the number of used classes, $n_{classes}$ and $n_{samples}$, as shown in Equation 4.1. This number $n_{samples}$ is later divided intro training (n_{train}) and test (n_{test}) samples in a ratio of 2/3 and 1/3, respectively.

$$n_{bag} = 11 \times n_{classes} \times n_{samples} \tag{4.1}$$

A point to be addressed is that the *bag* availability according to each Setup and *prop* used thus n_{bag} will not be the same. In case of Setups 1, 2 and 5, the *bag* availability for training samples to be selected will always be $n_{samples}$ per class as it regards an specific strata. On the other hand, for Setups 3, 4 and 6, as they use sets of $[prop, 100\%]$, defining accumulated strata, the *bag* availability will vary accordingly. As there are 11 strata per class in this methodology, the *bag* availability will vary according to Table 4.4.

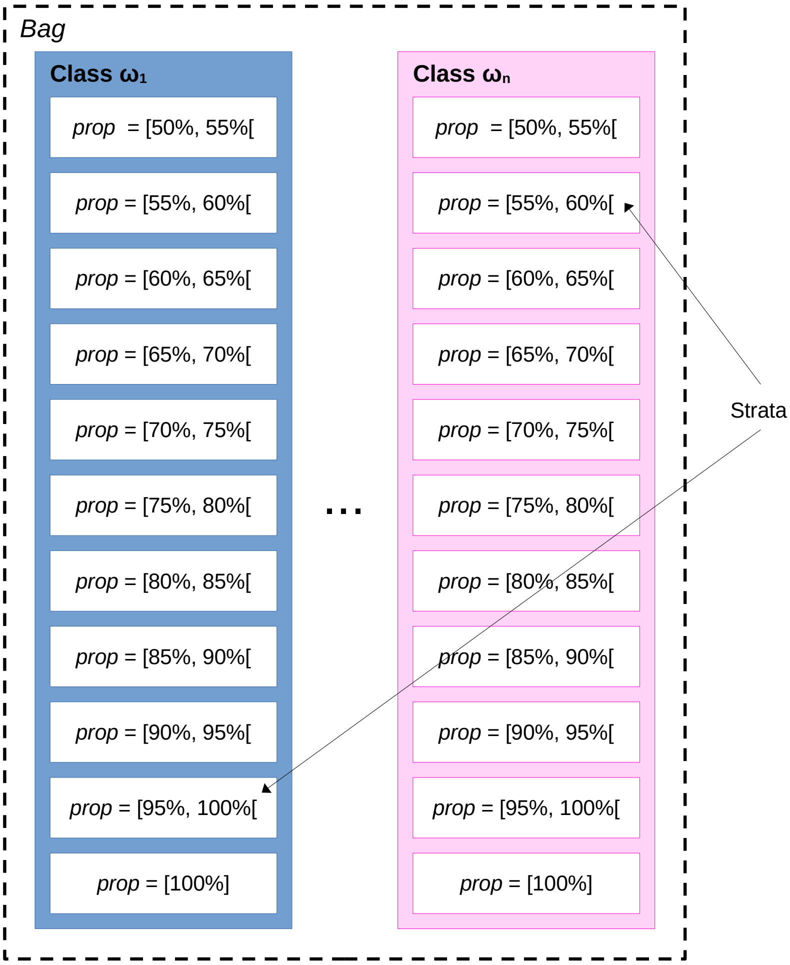
Table 4.4 - Reference Sample Selection - Bag size.

Proportion (*prop*)	Available samples in Bag
$\leq 50\%$	$11 \cdot n_{classes} \cdot n_{samples}$
$\leq 55\%$	$10 \cdot n_{classes} \cdot n_{samples}$
$\leq 60\%$	$09 \cdot n_{classes} \cdot n_{samples}$
...	...
$\leq 95\%$	$02 \cdot n_{classes} \cdot n_{samples}$
$= 100\%$	$01 \cdot n_{classes} \cdot n_{samples}$

The number of available strata per used *prop* for Setups 3, 4 and 6. Thus n_{bag} will be reduced according to it.

SOURCE: Author.

Figure 4.4 - Illustration of how the *bag* is stratified for the RSS process.



The illustration regards a population with n classes with $n \times 11$ strata, where each stratum is defined per class and per modal class proportion.

SOURCE: Author

After filtering the data in RSS Part I, defining the *bag* hence the determining candidates for reference data, the supervised classification using the Monte Carlo Simulation takes place. This process is repeated 100 times for each *prop*. Hence, for each Setup, there are $1,100$ image classifications. Two supervised non-parametric classifiers are used: K-Nearest Neighbours, with $k = 5$ (KNN-5) and Support Vector Machine One-Against-One (SVM-OAO). The classification is only conducted on the test samples for optimising the process.

The parameters used for each classifier during the Monte Carlo simulation were:

- **KNN-5:** considered an uniform weight for all neighbours with a leaf-size of 30; Euclidean distance;

- **SVM-OAO:** regularisation parameter: 1.0; the kernel being Radial Basis Function (RBF); gamma = *scale*; uses the shrinking heuristic and without iterations limit;

4.2.4 Reference Sample Selection - Thematic Accuracy and Completeness

At the end of each Monte Carlo repetition, the confusion matrix, overall accuracy and *kappa* index are computed in a way that, for each *prop* in a certain Setup, these average values and respective standard deviations are calculated and stored. Later, Completeness elements are computed in their averages, thus Omission Error (OE) and Commission Error (OE) are also computed, and their complements, which are Producer's Accuracy ($1 - OE$, PA) and User Accuracy ($1 - CE$, UA). After performing the classifications, charts are created so all objectives shown in Chapter 3 can be interpreted and discussed.

5 RESULTS AND DISCUSSION

The presented results refer to the two used Baseline approaches: Pixel-based (PIX) and Region-based (REG) and their results are presented separately in Sections 5.1 and 5.2, respectively.

For both situations, the results are presented as follows: (i) the Baseline Classifications results with the regions of interests (ROI), thematic accuracy, completeness and the baseline map followed by (ii) the extraction of candidate data in RSS Part I and (iii) the selection of reference data with the supervised image classification using Monte Carlo simulation in RSS Part II and (iv) spatial data quality regarding thematic accuracy and completeness separately for each used Setup.

We emphasise the amount of resulting charts and tables hence, for RSS results, only thematic accuracy is shown, though all results are presented in Appendix B.

5.1 Pixel-based baseline classification

5.1.1 Baseline classification

For the pixel-based baseline classification, the analyst manually selected the reference data, assuming approximately the same amount of samples for class, with their values shown in Table 5.1. The spatial distribution of these samples is presented in Figure 5.1. Also, these samples were randomly divided into 70% for training and 30% for test (spatial data quality). Therefore, the sampling design is the stratified random sampling.

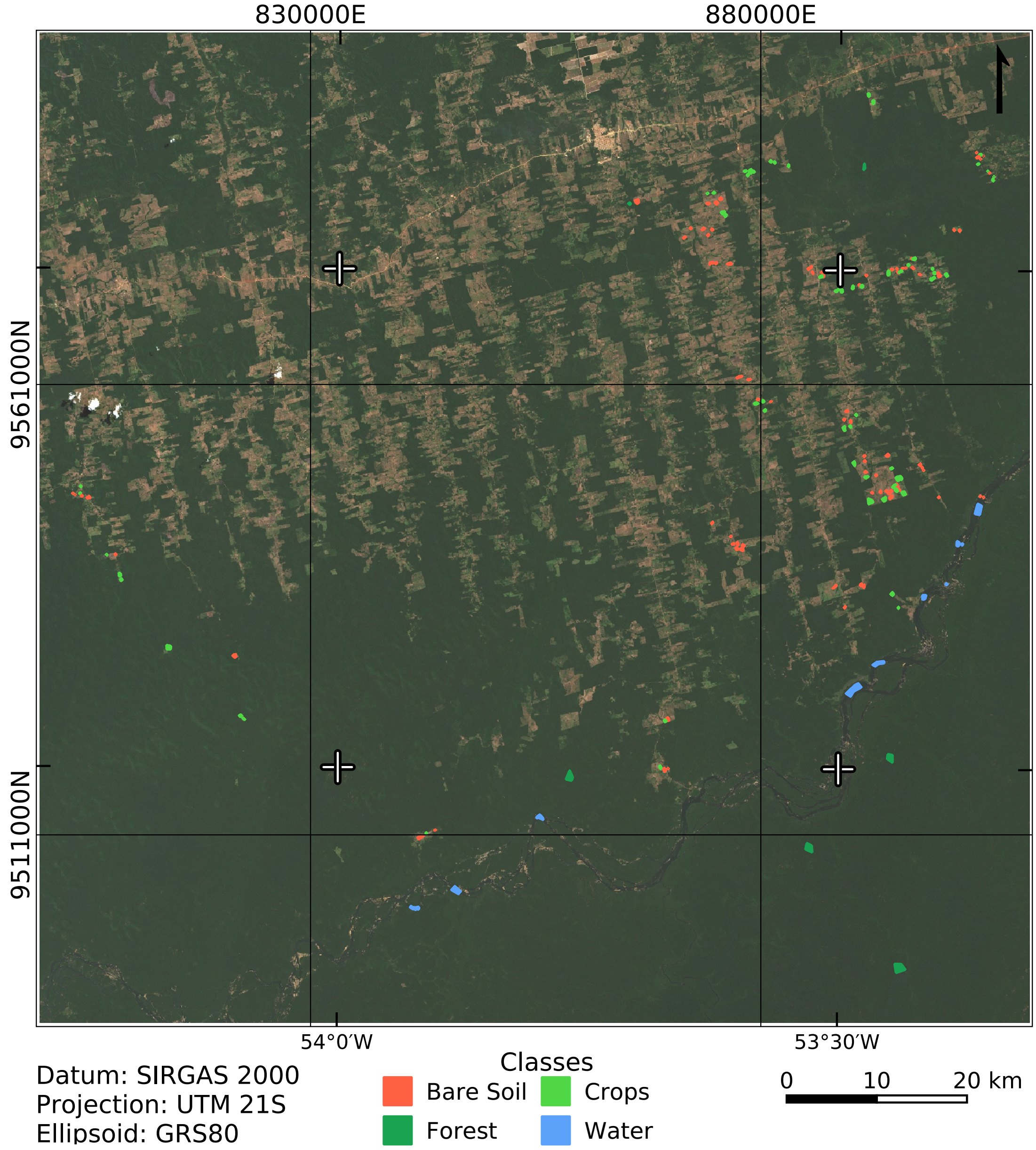
This approach used the Random Forest classifier (RF) and, for this classification, the OOB score was 97.84%, with an order of importance and their respective percentage of importance shown in Table 5.2. The order of importance showed that the Green band (B03) was the most relevant feature for separating the classes followed by NIR (B08), Blue (B02) and Red (B04).

Table 5.1 - Pixel-based First Classification Total number of samples per class.

Class	Water	Forest	Crops	Bare Soil
Total number of samples	26,576	20,226	28,429	20,323

SOURCE: Author.

Figure 5.1 - Pixel-based Baseline Classification using Random Forest - Reference samples.



Selected samples presented over a Sentinel-2A image TCI band from 9 August 2020. The polygons have been augmented for better visualisation. This Figure in Black and White is presented in Figure C.2, page 198.

SOURCE: ESA (2015a).

Table 5.2 - Pixel-based Baseline Classification using Random Forest - Order of Importance.

Band	Order of Importance (%)
Green (B03)	35.36
NIR (B08)	31.78
Blue (B02)	16.72
Red (B04)	16.19

SOURCE: Author.

When it comes to Spatial Data Quality, the confusion matrix generated from RF is presented in Table 5.3. From this table, Thematic Accuracy was computed, where overall accuracy was 98.096% and *kappa* index was 0.974, meeting Condition IV. As we can see, OOB Score and OA present similar results, as it is expected due to RF characteristics that may put the need of presenting a thematic accuracy in question. In this case, as there are two approaches for baseline classification, the thematic accuracy is used for comparisons and for being in accordance with the Brazilian Standards of Spatial Data Quality (DSG, 2016). This high value indicates that there are almost no errors, and this is endorsed by the confusion matrix.

Additionally, the completeness elements were computed and are presented in Table 5.4. The least accurate class regarding user accuracy (UA) is Crops that confuses with Forest and Bare Soil, while for Producer Accuracy, Forest is the least accurate one, confusing with Crops. Nonetheless, the total percentage of errors is considerably small, hence may be neglected.

Therefore, at the end of this step, Conditions I, II, III and IV are met. The next step is to determine the quality of LR cells with modal class proportion (*prop*), which is part of the Reference Sample Selection (RSS) Part I.

Table 5.4 - Pixel Based First Classification using Random Forest - Completeness.

Class	User Accuracy (%)	Producer Accuracy (%)
Water	100.000	100.000
Forest	97.198	96.414
Crops	96.244	97.047
Bare Soil	98.712	98.375

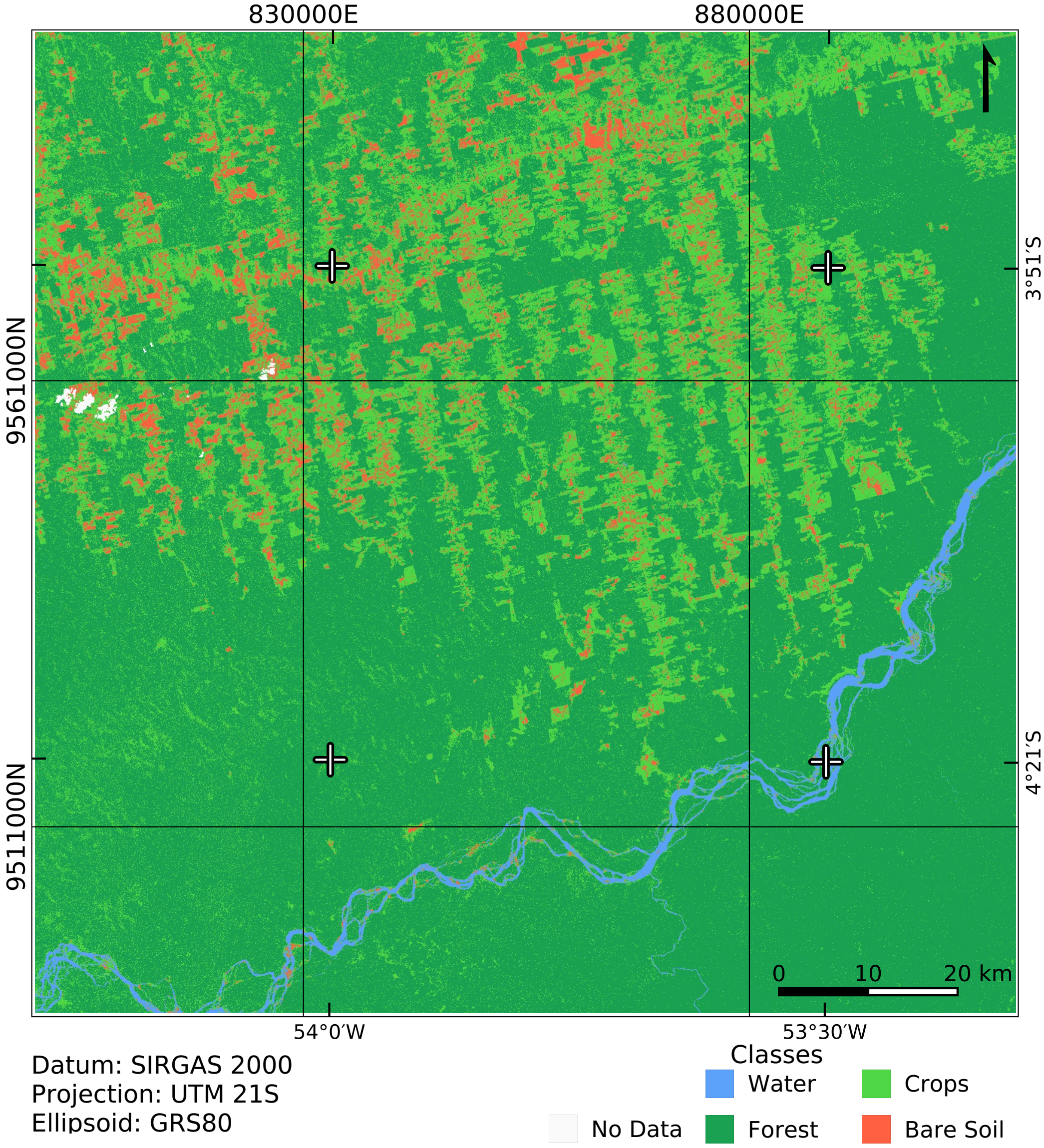
SOURCE: Author.

Table 5.3 - Pixel-based Baseline Classification using Random Forest - Confusion Matrix.

	Class	Water	Forest	Crops	Bare Soil	**Total**
		Predicted				
Truth	Water	**8035**	0	0	0	**8035**
	Forest	0	**5862**	169	0	**6031**
	Crops	0	218	**8149**	100	**8467**
	Bre Soil	0	0	79	**6055**	**6134**
	Total	**8035**	**6080**	**8397**	**6155**	**28667**

SOURCE: Author.

Figure 5.2 - Pixel based First Classification using Random Forest - classified image.



White pixels correspond to NoData values. This image in Black and White is presented in Figure C.3, page 199

SOURCE: Author.

5.1.2 Reference Sample Selection part I - filtering candidate samples

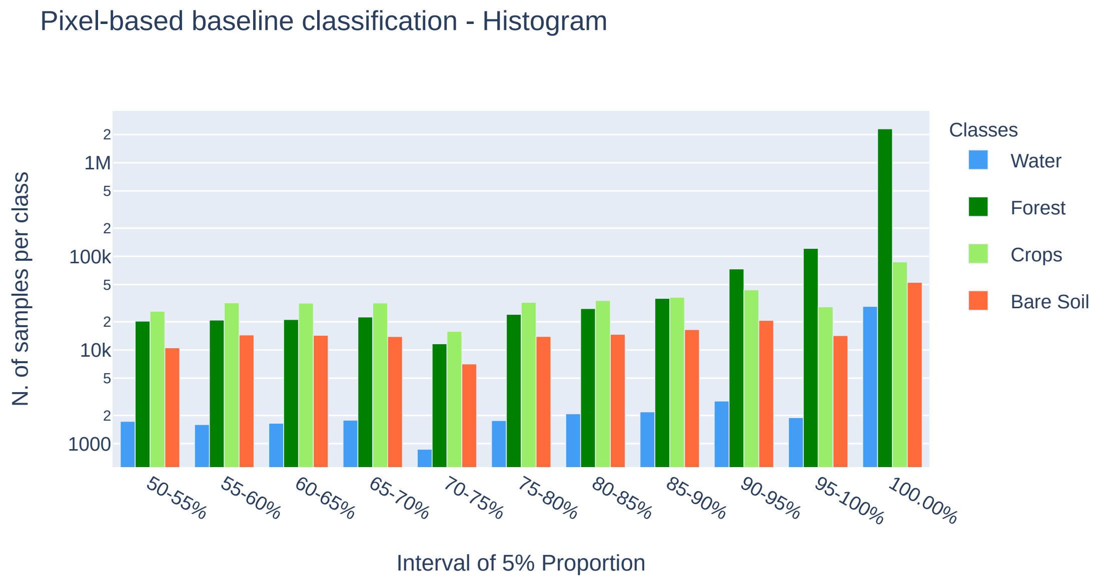
After the baseline image is generated, the first step of the RSS approach is to acquire the modal class and its frequency in each LR cell considering a minimum *prop* of 50% to meet Condition V (chapter 3). More than half the image was selected as candidate reference data, in a total of $3,320,844$ cells. Three different ways of presenting these candidate samples are used: Histogram, Pie Chart, Table and Frequency Map.

The first ones is the Histogram (Figure 5.3) of (i) accumulated number of samples per class in the upper chart and (ii) number of samples per stratum in the lower chart. This Histogram is in logarithm scale due to the great amount of candidate cells for Forest class with $prop = 100\%$. This figure can aid the interpretation of class representativity: the majority of Forest and Water candidate samples belongs to $prop = 100\%$, which is not true for the other classes. Crops and Bare Soil visually present a more balanced distribution.

The second way are pie charts showing the distribution of candidate data for each stratum (Figure 5.11). These charts illustrate that Water and Forest have their great majority of data on $prop = [100\%]$ whilst this is not true for the other classes, implying the importance that the study of class' representativity, as shown in Table 5.5.

The third way is Table 5.6, which created the former figure. This table also presents the percentage of candidate samples per stratum in a certain class. When analysing the number of candidate samples in classes with $prop = 100\%$, Water and Forest have values in 61.37% and 85.87%, respectively. For this specific study area, these values strongly indicate that these two classes, purer samples are representative of them. Bare Soil and Crops have their percentages of samples for pure samples at 21.81% and 27.82%. Even though most of the samples in these classes are in $prop = 100\%$, the majority does not achieve half of the candidate data. Therefore, these classes' representatives are more distributed, leading us to the conclusion that a variety of sample quality may be better when choosing samples for theses classes than homogeneous samples.

Figure 5.3 - Pixel-based Baseline Classification using Random Forest - histogram.



Number of cells presented on the superior image in accumulated samples per modal class proportion on set $[prop, 100\%]$ and in the inferior image with samples per interval of modal class proportion $[prop, prop + 5\%[$.

SOURCE: Author.

Additionally, from the Table 5.6, we can analyse the class distribution, as presented in Table 5.5. For PIX, the averages grouped values varied from 83% up to 97%. Forest class showed the lowest *std* indicating a low distribution of the class; the negative skewness indicates the central mass on the right side of the histogram, as can be perceived in Figure 5.3; the positive excess kurtosis shows a leptokurtic distribution in a shape of an inverted "V". Crops and Bare Soil presented similar values of distribution; their skewness are closer to zero, indicating slightly more symmetric tail compared to the other classes; their excess kurtosis are negative, indicating platykurtic curve (i.e. a more flattened distribution). Water class presented negative skewness indicating that the mass centre is more to the right side of the histogram and negative excess kurtosis, indicating a more platykurtic curve than the other classes. From these information, we can study each class representativiness with respect to the sample quality. Bare Soil, Crops and Water are more represented by mixed quality sample, while Forest are best represented by pure samples.

Table 5.5 - Pixel-based Baseline Classification using Random Forests - histogram distribution statistics.

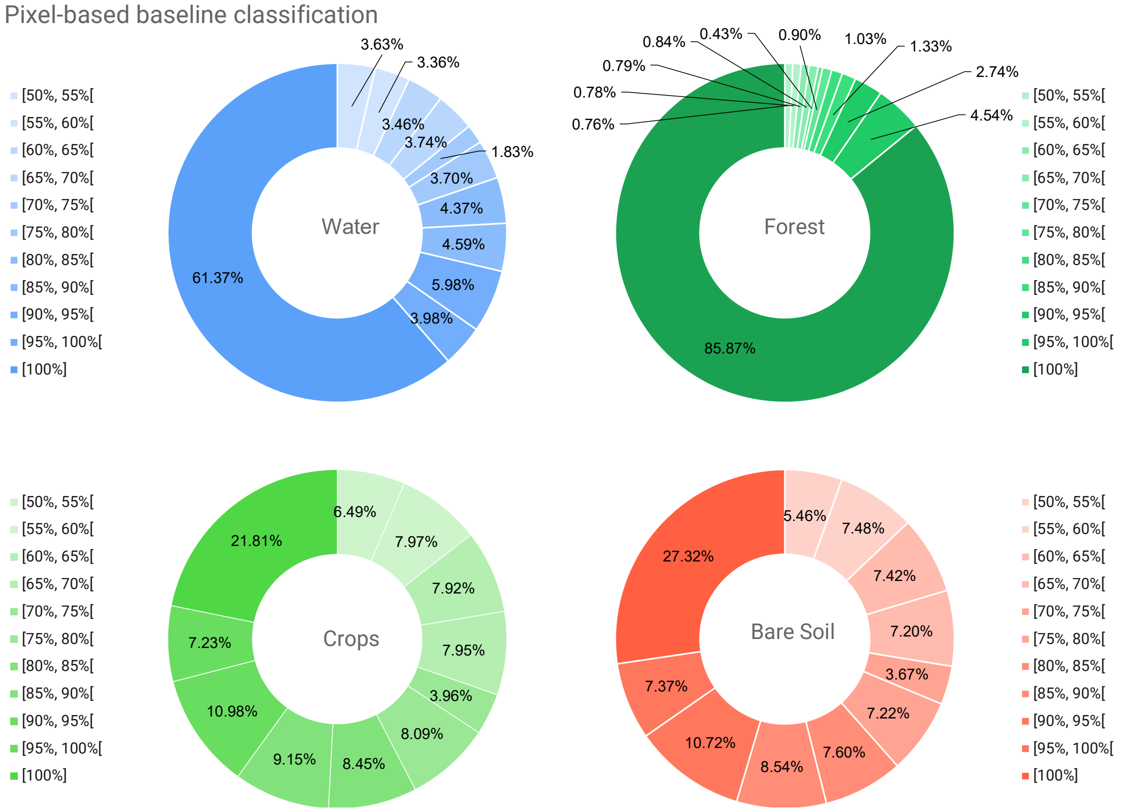
Class	**Grouped mean** (%)	**Grouped *std*** (%)	**Skewness**	**Excess kurtosis**
Water	91.116	14.438	-1.485	-2.591
Forest	97.754	7.807	-4.156	17.312
Crops	81.440	16.093	-0.382	-1.246
Bare Soil	83.084	16.125	-0.513	-1.161

SOURCE: Author.

Finally, a spatial representation of the proportion is shown in Figure 5.5. The spatial distribution varies from lighter colours ($prop = 50\%$) to darker colours ($prop = 100\%$). This image particularly aggregated pair of stratum to ease the image visualisation hence it used the interval $[prop, prop + 10\%[$ and $prop = 100\%$. This map can delineate border LR cells, showing the fishbone pattern of deforestation in this area as well as river margins. Another point is that the noises from the baseline classification reflect directly to the RSS process and can be seen in this figure. This fact highlights how the quality of the baseline classification affects the RSS process.

All these representations infer that, as expected, border samples would present a lesser quality ($prop$), and, for this specific scenario, Water and Forest classes showed a better representation with $prop = 100\%$ whilst Crops and Bare Soil are not well represented by it. Thus, selecting higher quality, or pure, samples may not always be the best choice to represent all the classes. The questions remaining are: to what extent this statement is valid? And for which classes? Subsequent analyses of RSS approach may aid to some conclusions.

Figure 5.4 - Pixel-based Baseline Classification using Random Forest - pie charts of distributed data per class.



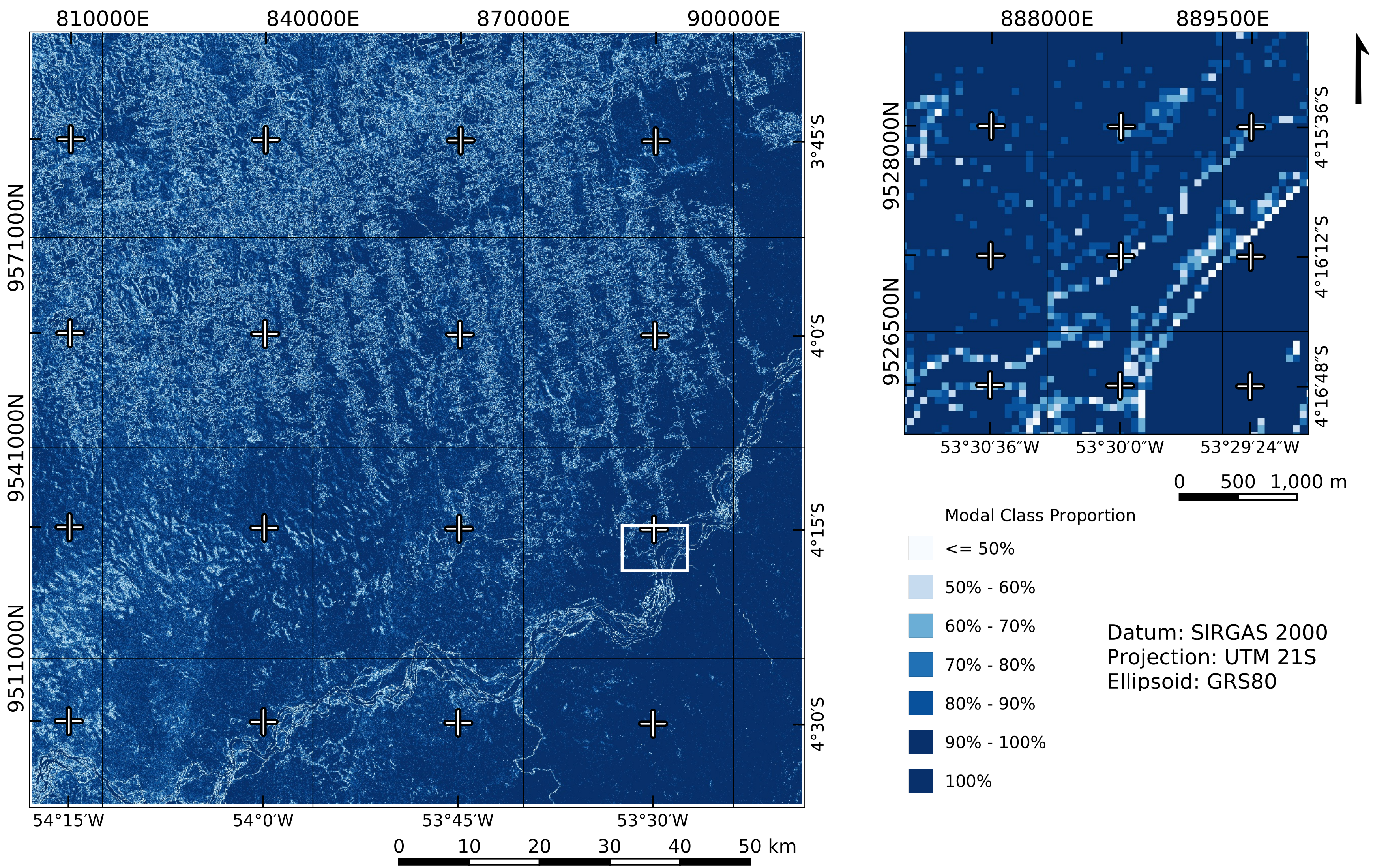
SOURCE: Author.

Table 5.6 - Pixel-based Baseline Classification - number of samples per class and their respective percentage.

Proportion		**Number of samples**									
		Water		**Forest**		**Crops**		**Bare Soil**		**Total**	
		N.	(%)	N.	(%)	N.	(%)	N.	(%)	N.	(%)
50%	55%	1,727	3.63	20,320	0.76	25,898	6.49	10,552	5.46	58,497	1.76
55%	60%	1,597	3.36	20,861	0.78	31,840	7.97	14,466	7.48	68,764	2.07
60%	65%	1,649	3.46	21,150	0.79	31,640	7.92	14,346	7.42	68,785	2.07
65%	70%	1,778	3.74	22,497	0.84	31,739	7.95	13,917	7.20	69,931	2.11
70%	75%	870	1.83	11,647	0.43	15,797	3.96	7,093	3.67	35,407	1.07
75%	80%	1,761	3.70	24,010	0.90	32,282	8.09	13,947	7.22	72,000	2.17
80%	85%	2,082	4.37	27,685	1.03	33,753	8.45	14,693	7.60	78,213	2.36
85%	90%	2,183	4.59	35,568	1.33	36,538	9.15	16,511	8.54	90,800	2.73
90%	95%	2,846	5.98	73,362	2.74	43,836	10.98	20,713	10.72	140,757	4.24
95%	100%	1,893	3.98	121,649	4.54	28,855	7.23	14,250	7.37	166,647	5.02
100%		29,211	61.37	2,301,935	85.87	87,101	21.81	52,796	27.32	2,471,043	74.41
Total		**47,597**	100.00	**2,680,684**	100.00	**399,279**	100.00	**193,284**	100.00	**3,320,844**	100.00

SOURCE: Author.

Figure 5.5 - Pixel-based Baseline Classification using Random Forest - frequency map.



SOURCE: Author.

5.1.3 Reference Sample Selection part II - selecting reference data and image classification

Once the modal class proportion is studied in RSS Part I, the next step is defining the *bag* size. From Table 5.6, the least occurring stratum is Water in the set $[70\%, 75\%[$, which is 870. In order to compare these outcomes with those from the region-based baseline, we set $n_{samples}$ to 750. Hence, for each stratum - determined by intervals of 5% proportion per class - there will be 750 samples per stratum in the *bag*.

Consequently, the *bag* size is $33,000$ ($750 \times 11_{intervals} \times 4_{classes}$). For the Monte Carlo simulation, the split of the reference data was 500 for training and 250 for test samples per class, in a total of $2,000$ training samples and $1,000$ test samples for each image classification. Although the number of total reference data should be greater for an image this size, this could not be achieved due to the number of samples for Water class in all strata so the analysis could be affected.

For each Setup, the Monte Carlo Simulation with 100 repetitions run in an average of of three minutes and are presented in Spatial Data Quality section.

5.1.4 Reference Sample Selection - Spatial Data Quality

RSS Part II with Monte Carlo Simulation considered the six different Setups, using the *bag* to determine n_{train} and n_{test}, as presented herein. This simulation run 100 repetitions per modal class proportion or interval of proportion, in a total of $1,100$ repetitions per Setup. The classifiers used were SVM-OAO and KNN-5, as presented in Section 4.2.3. Additionally, thematic accuracy and completeness were computed, considering their respective means and standard deviations.

These two Spatial Data Quality components were treated separately so we can focus firstly on thematic accuracy (Section 5.1.4.1), and later on completeness (Section 5.1.4.2), mainly observing how the used classes respond to the quality variation of reference data.

Additionally, the tables with exact information regarding Spatial data Quality are presented in Appendix B.1.2. These tables present values of *kappa* index and overall accuracy (OA) with their respective *std*. They also present User Accuracy (UA) and Producer Accuracy (PA) in percentage values for each class.

The confusion matrices for all setups and intervals are presented from page 151 on, in Appendix B.1.3.

5.1.4.1 Thematic accuracy

The results for thematic accuracy were separated in two groups of discussion. The first group concerns Setups 3, 4 and 6 (Figure 5.6) and the second group concerns Setups 1, 2 and 5 (Figure 5.7). A reason for this division lies on how the training samples were set, as the former 3 Setups could be performed by an analyst. Also, the analyses take into account three points to answer the scientific question: (i) how the used classifier responds to the variation of reference data quality; (ii) how thematic accuracy is affected by the convergence of reference data quality and (iii) how the thematic accuracy is affected by training and test samples separately. Addisionally, these results are also presented in tabular results in Table 5.7 and Table 5.8.

Firstly considering the first group (Figure 5.6), containing Setups 3, 4 and 6, with training samples in set $[prop, 100\%]$. This set evolution can be comparable to the evolution in experience of an analyst. In this group, all *kappa* values were greater than 0.75. Setup 3, with test samples in set $[50\%, 100\%]$ had its *kappa* values decreasing as *prop* increased; this setup shows how training and test samples react as their quality diverge from each other, once the quality of test samples was invariant and training samples increased their quality. In this case, KNN-5 seemed to be more sensitive to this divergence as the *kappa* decreased and the *std* increased. For SVM-OAO, the *kappa* values stabilised, showing a more flexible feature space separation which is inherent of this classifier.

Moreover, for both classifiers, KNN-5 and SVM-OAO, Setup 4 had its results close to 1.00; this Setup has test samples in $[100\%]$. So this Setup results are very high independently of the quality of training samples for both classifiers, which implies that the quality of test samples may be more relevant than the quality of training samples though it cannot be yet confirmed.

Finally, for Setup 6, with test in $[prop, 100\%]$, the qualities of both test and training data increase together. Even though there is a growth on the *kappa*, it starts at ≈ 0.85 (SVM-OAO), which may be explained by the classes' divergent spectral responses. This shows that, as the analyst experience increase, so does the resulting thematic accuracy, reaching up to almost $kappa = 1.00$. Probably, if there were more similar classes used and if it was not a controlled situation, this initial *kappa* would be smaller. Despite of that, considering the classifiers sensibility, SVM-OAO seemed to be slightly more sensitive to the quality of both training and test data quality, as initial *kappa* (i.e. $prop = 50\%$) was inferior for this classifier.

Moving on to the analysis of the second group (Figure 5.7), formed by Setups 1, 2 and 5, their training set are in $[prop, prop + 5\%[$, these setups are completely theoretical and were created to aid discussions over each reference data quality.

Beginning with Setup 1, with test set in $[50\%, 100\%]$, it is clear that the qualities of test and training samples are always different, so we cannot infer over convergence of training and test data. However, a fact to be pointed out is that both classifiers responded differently to this Setup: KNN-5 decreased while SVM-OAO stabilised when $prop \geq 65\%$. A likely reason for that is on the way each classifier works to separate the feature space into classes. For SVM-OAO, the creation of hyperplanes can already be satisfactory with training data in a quality around 65%. Controversially, for KNN-5 the resulting *kappa* decreased, inferring that this classifier relies mostly on the quality of the training data, if the test data has mixed quality. Therefore, SVM-OAO seems to separate better theses classes when modelling the prediction model than KNN-5.

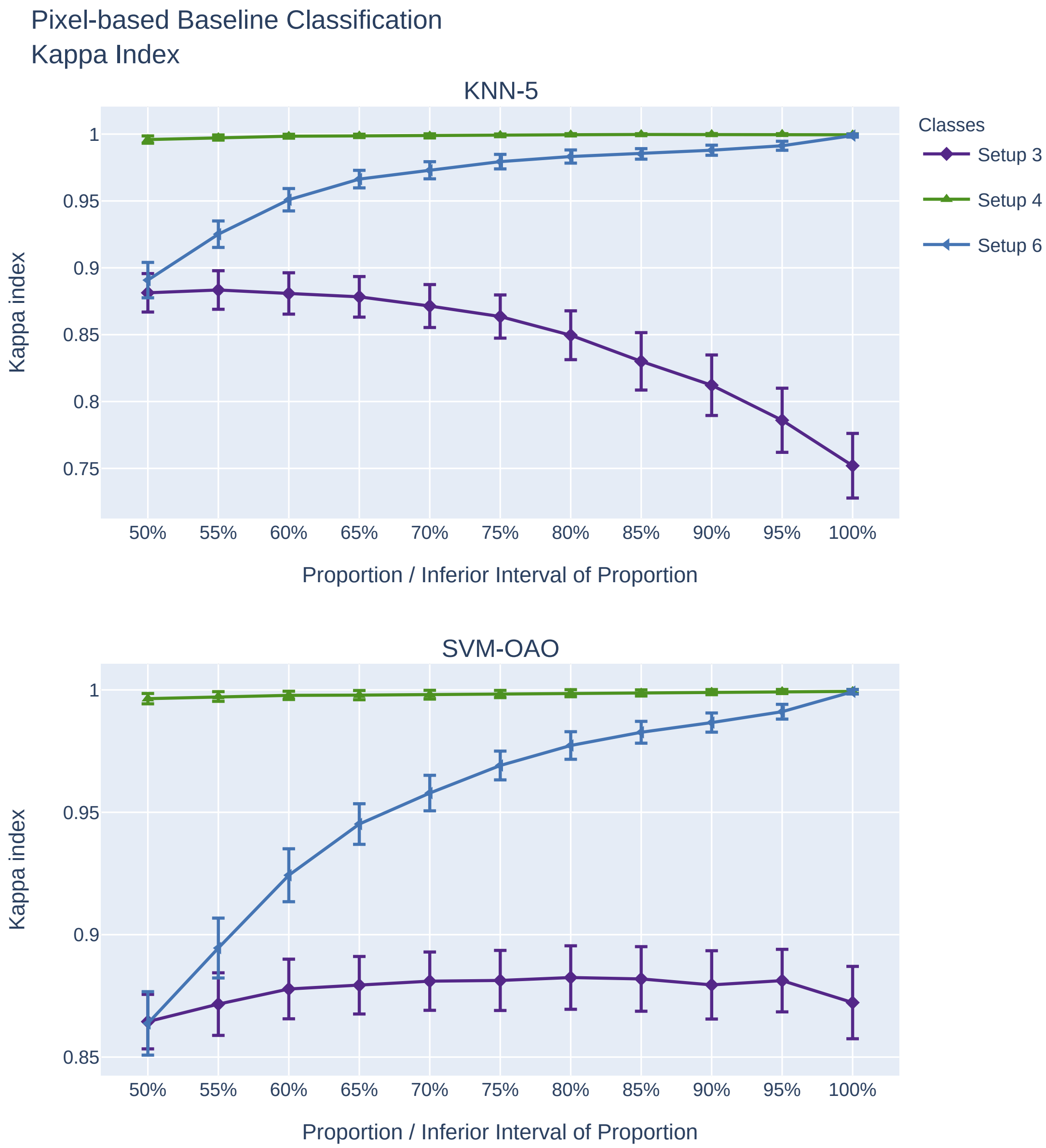
Setup 2, with test set in $[100\%]$ showed that, for both classifiers, independently of the quality of training data, when above 60%, the *kappa* index reaches values close to 1.00. Still, when analysing other *prop* values, all of them presented relatively high *kappa* (≥ 0.80), which restated the conclusion for Setup 4: the thematic accuracy is most affected by the quality of test data than training data. Hence, the outcomes for this Setup emphasises that the quality of training data, as long as it is above 60%, is almost irrelevant for defining a high thematic accuracy when compared to a pure test data.

The last Setup to be studied is Setup 5 (test in $[prop, 100\%]$). The test set can be compared to an analyst experience while we analyse strata of quality. For both classifiers there was an increase though *kappa* stabilised in KNN-5 and kept growing in SVM-OAO. We can observe that, once $prop \geq 70\%$, the *kappa* values are above 0.95. This indicates that sample quality above 70%, for training and test data, may already present high thematic accuracy.

Focusing on the training samples, we can study pairs of Setups: 1 and 3, 2 and 4, 5 and 6. All of them showed a tendency to converge as *prop* grew indicating that the higher the *prop* the lower the influence of training samples in the thematic accuracy. For all the three pairs, when $prop \leq 75\%$, training samples tend to affect more this metric. This shows that the thematic accuracy is not strongly sensitive to training samples.

Therefore, when analysing thematic accuracy for Pixel-based Baseline classification and for the four classes studied (Water, Forest, Crops and Bare Soil), we can observe that (i) SVM-OAO may separate better the feature spaces, hence it differentiates better the classes during the prediction model; (ii) the quality of test data affects more significantly the thematic accuracy outcomes; (iii) when there is difference in the quality of test/training data, the classifier may respond poorly to it. From these observations, few questions came up: to what extent can we trust the thematic accuracy as a metric to determine a good image classification? Is it possible that the thematic accuracy can be biased by the quality of the test data? How can we relate the class representativity with a high *kappa* index? Obviously, this is a simulated data with very separable classes and this may overestimate the outcomes. Still, when analysing group 2 (Setup 1, 2 and 5), it is clear that the quality of reference data affects the image classification.

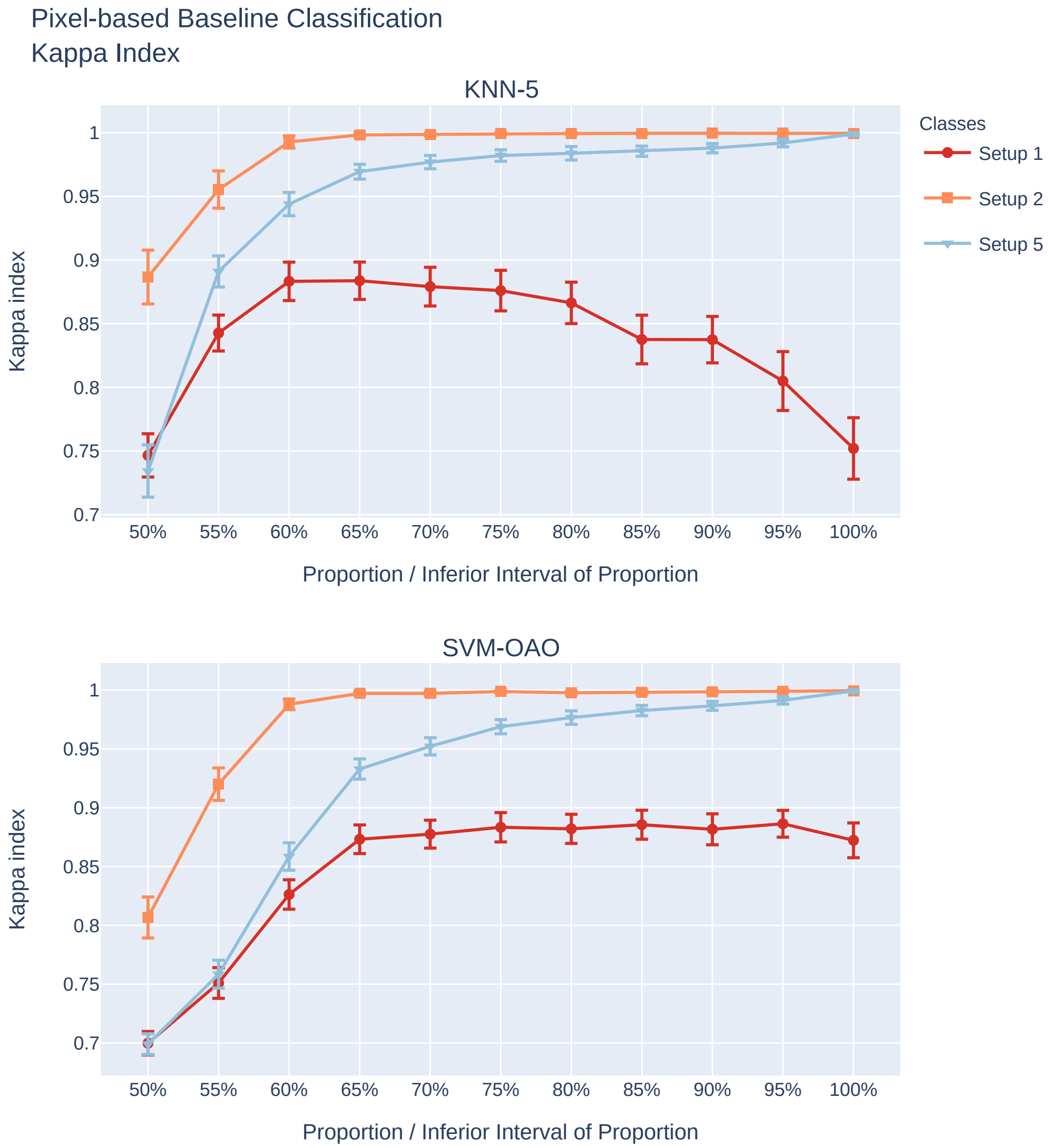
Figure 5.6 - Pixel-based Baseline Classification using Random Forest - Thematic Accuracy with error bars for Setups 3, 4 and 6.



Where Setup 3 corresponds to train: $[prop, 100\%]$ / test:$[50\%, 100\%]$; Setup 4 corresponds to train: $[prop, 100\%]$ / test:$[100\%]$; Setup 6 corresponds to train: $[prop, 100\%]$ / test: $[prop, 100\%]$. More detailed results are presented in Appendix B.1.

SOURCE: Author.

Figure 5.7 - Pixel-based Baseline Classification using Random Forest - Thematic Accuracy with error bars for Setups 1, 2 and 5.



Where Setup 1 corresponds to train: $[prop, prop + 5\%[$ / test:$[50\%, 100\%]$; Setup 2 corresponds to train: $[prop, prop + 5\%[$ / test:$[100\%]$; Setup 5 corresponds to train: $[prop, prop+5\%[$ / test: $[prop, 100\%]$. More detailed results are presented in Appendix B.1.

SOURCE: Author.

Table 5.7 - Pixel-based Baseline Classification - Thematic accuracy and standard deviation for KNN-5.

prop	**Setup 1**		**Setup 2**		**Setup 3**		**Setup 4**		**Setup 5**		**Setup 6**	
	kappa	std	*kappa*	std	*kappa*	std	*kappa*	std	*kappa*	std	*kappa*	std
50%	0.747	0.017	0.887	0.021	0.881	0.014	0.996	0.003	0.734	0.021	0.891	0.013
55%	0.843	0.014	0.955	0.015	0.883	0.014	0.997	0.002	0.891	0.012	0.925	0.010
60%	0.883	0.015	0.993	0.005	0.881	0.015	0.998	0.001	0.944	0.009	0.951	0.008
65%	0.884	0.015	0.998	0.002	0.878	0.015	0.999	0.001	0.969	0.006	0.966	0.007
70%	0.879	0.015	0.999	0.001	0.871	0.016	0.999	0.002	0.977	0.005	0.973	0.006
75%	0.876	0.016	0.999	0.001	0.864	0.016	0.999	0.001	0.982	0.005	0.979	0.005
80%	0.866	0.016	0.999	0.001	0.850	0.018	1.000	0.001	0.984	0.005	0.983	0.005
85%	0.838	0.019	0.999	0.001	0.830	0.021	1.000	0.001	0.985	0.004	0.985	0.004
90%	0.837	0.018	1.000	0.001	0.812	0.023	1.000	0.001	0.988	0.004	0.988	0.004
95%	0.805	0.023	1.000	0.001	0.786	0.024	1.000	0.000	0.992	0.003	0.991	0.003
100%	0.752	0.024	0.999	0.001	0.752	0.024	0.999	0.001	0.999	0.001	0.999	0.001

SOURCE: Author.

Table 5.8 - Pixel-based Baseline Classification - Thematic accuracy and standard deviation for SVM-OAO.

prop	**Setup 1**		**Setup 2**		**Setup 3**		**Setup 4**		**Setup 5**		**Setup 6**	
	kappa	std	*kappa*	std	*kappa*	std	*kappa*	std	*kappa*	std	*kappa*	std
50%	0.700	0.010	0.807	0.017	0.864	0.011	0.996	0.002	0.699	0.009	0.864	0.013
55%	0.751	0.013	0.920	0.014	0.872	0.013	0.997	0.002	0.758	0.012	0.895	0.012
60%	0.826	0.013	0.988	0.005	0.878	0.012	0.998	0.002	0.858	0.012	0.924	0.011
65%	0.873	0.012	0.997	0.002	0.879	0.012	0.998	0.002	0.933	0.009	0.945	0.008
70%	0.878	0.012	0.997	0.002	0.881	0.012	0.998	0.002	0.952	0.007	0.958	0.007
75%	0.883	0.012	0.999	0.001	0.881	0.012	0.998	0.001	0.969	0.006	0.969	0.006
80%	0.882	0.012	0.998	0.002	0.882	0.013	0.999	0.001	0.977	0.006	0.977	0.006
85%	0.886	0.012	0.998	0.002	0.882	0.013	0.999	0.001	0.983	0.004	0.983	0.004
90%	0.882	0.013	0.998	0.001	0.879	0.014	0.999	0.001	0.987	0.004	0.987	0.004
95%	0.886	0.011	0.999	0.001	0.881	0.013	0.999	0.001	0.991	0.003	0.991	0.003
100%	0.872	0.015	0.999	0.001	0.872	0.015	0.999	0.001	0.999	0.001	0.999	0.001

SOURCE: Author.

5.1.4.2 Completeness

Completeness refers to a over or under-completeness analysis of the image classification. This study will analyse it separately from thematic accuracy in order to compare the classes performance for each Setup. The charts for this section are presented in Appendix B.1 and are presented per Setup. Regarding the tables, they are presented in Appendix B.1.2 with the confusion matrices from page 152.

5.1.4.2.1 Setup 1

These Setup studies intervals of increasingly reference data quality (training in set $[prop, prop{+}5\%[$) against a fixed mixed data quality ($[50\%, 100\%]$). For Pixel-Based Baseline Classification, the images with results are shown in Figure B.1 and Figure B.2. Also, an illustration of this Setup classes' pattern is presented in Figure A.1.

Forest presented decreasing user accuracy for KNN-5 and SVM-OAO, though SVM-OAO had a lower variation (max 99.60% and min 94.63%) than KNN-5 (max 93.38% and min 75.96%). This indicates that the usage of a certain classifier may have some effect upon the the resulting completeness. When it comes to producer accuracy, there was also a variation between the two classifiers, as KNN-5 rose from 75.78% to 98.18% while SVM-OAO rose from 68.41% to 88.33%, showing a similar range. This indicates that, generally, producer accuracy in Forest class is more sensitive to quality variation than user accuracy. This means that, for the Forest scenario, the higher the quality we use for training the classifier, the more classes will be mistakenly assigned to it. A probable reason for that is how the division of the feature space given the quality of reference data for Setup 1, which is illustrated in Figure A.1. This Setup increases the quality of training data, hence the classifiers tend to separate better the feature spaces. However, the test samples are always mixed, showing that SVM-OAO dealt better with this situation than KNN-5 mainly because KNN-5 is distance dependent.

Crops presented a high range of user and producer accuracies for both classifiers. Regarding the user accuracy, it varied from 48.36% to 94.52% in KNN-5 and from 13.95% to 82.12%. This shows that, unlike Forest, the quality of training data has a relevant effect on Crops UA, when test data is mixed and SVM-OAO seemed to be more sensitive to that. On the other hand, when analysing producer accuracy, the accuracy decreased for both classifiers: KNN-5 started its PA at 70.01%, raised to 86.07% ($prop = 65\%$) then fell to 68.39%; while SVM-OAO started at 96.30%, had its maximum value at 96.61% ($prop = 55\%$) then decreased to 81.78%. Hence, in general, the greater the *prop*, the greater the absence of data as Crops is. This class presented the highest variation and, when compared to other classes, had a different pattern of UA and PA. A likely reason for that is that Crops seems to be dispersed in the feature space. As KNN-5 and SVM-OAO use distances in the feature space to delineate classes divisions, then when a class presents dispersed characteristics, it does not seem to be well divided with pure training samples.

In general, Bare Soil varied differently in both classifiers UA and PA. For KNN-5 UA, this class started at 82.41% increased to 92.94% ($prop = 65\%$) then fell to 79.19%. On the other hand, in SVM-OAO UA, it varied from 99.87% to 91.21%. This showed that as *prop* increased, the less the reliability on this class was. although SVM-OAO seemed to have dealt better with this quality variation. Regarding PA, for KNN-5, it started at 78.70%, rose to 91.65% ($prop = 70\%$) then fell to 73.03%. In SVM-OAO PA, it started at 70.41% then rose up to 92.82%. This elucidates the discrepancy in which the classifiers affect completeness: KNN-5 is more sensitive to the class dispersion in the feature space than SVM-OAO.

For this Setup in almost all situations the accuracy values were above 75%, showing that Water class seems to be more homogeneous and well divided in the feature space, independently on the used training quality set. Water PA values converged to 100% for both classifiers and had some decreased in UA values. KNN-5 UA fell from 99.82% to 75.94% and SVM-OAO UA fell from 96.53% to 92.23%. This emphasises that the higher the *prop* the lower the class reliability is when test samples are mixed.

Then, from this Setup we can conclude that the classifiers affect the completeness, varying drastically the feature space separability. Also, considering mixed test samples, purer training samples did not present the best outcomes.

5.1.4.2.2 Setup 2

Setup 2, with training samples in $[prop, prop{+}5\%[$ and test samples in $[100\%]$, has its graphs shown for Pixel-Based Baseline Classification, in Figure B.3 and Figure B.4. Also, an illustration of this Setup classes' pattern is presented in Figure A.2.

This Setup presented KNN-5 and SVM-OAO UA converging to 100% when $prop \geq 65\%$. For KNN-5 UA, Water and Forest were always close to 100% while Bare Soil and Crops had greater variation. For SVM-OAO UA, Crops was the only class that diverged from the remaining, with its lowest UA of 42.87%.

Consequently, as thematic accuracy concluded, the quality of test samples interfere more dramatically in the accuracy assessment if the quality of training samples is $\geq 65\%$, independently of how the class is distributed in the feature space.

5.1.4.2.3 Setup 3

Setup 3, with training samples in $[prop, 100\%]$ and test samples in $[50\%, 100\%]$, has its graphs shown for Pixel-Based Baseline Classification, in Figure B.5 and Figure B.6. Also, an illustration of this Setup classes' pattern in presented in Figure A.3.

Similarly to Setup 1, Forest presented different accuracy pattern for UA and PA: it fell in the former while it rose in the latter. Regarding KNN-5 UA, it fell from 90.60% to 75.95% and for SVM-OAO UA, it fell from 97.81% to 96.12%. When analysing PA, for KNN-5 it rose from 91.78% to 98.18% and for SVM-OAO the rise was from 82.40% to 88.33%. These outcomes confirm the interference of the used classifier to divide the feature space and it also shows that the likelihood of having samples from other classes increased as *prop* increased, showing that other classes dispersion in the feature space affects the UA of more homogeneous classes. This does not occur for PA, as *prop* increases together with PA, pointing out the homogeneity of this class.

Considering UA, for KNN-5 it rose from 82.62% to 94.52% and for SVM-OAO, it grew from 67.60% to 82.12%. This pattern did not occur for PA, since it fell from 83.95% to 68.39% (KNN-5) and from 92.28% to 81.78% (SVM-OAO). These values, as expected, are similar to Setup 1 and shows that for more dispersed classes in the feature space, when trained in better sample qualities, these classes tend to be more mistakenly assigned to other classes.

Regarding user accuracy, this class showed similar results to Forest, presenting a falling curve for both classifiers. KNN-5 fell from 90.04% to 73.03% while SVM-OAO fell from 97.47% to 91.21%. For producer accuracy, it fell in KNN-5 from 91.90% to 79.19% and in SVM-OAO from 88.43% to 92.82%, which reiterates Setup 1 results.

Water class presented its greatest range in KNN-5 UA, falling from 99.27% to 75.94%. This indicates similar results to Forest regarding the interference of heterogeneous classes for well defined classes, like Water, in KNN. This interference exists though in a smaller scale for SVM-OAO, where the UA fall varied from 96.46% to 92.23%. Regarding producer accuracy, as this is a homogeneous class, both classifiers seem to converge to $\approx 100\%$: 98.48% to 99.92% (KNN-5) and 98.61% to 99.71% (SVM-OAO).

In general, results from this Setup are similar to the outcomes from Setup 1, showing the relevance of lesser quality data for the study of spatial data quality. Clearly, heterogeneity of classes interfere drastically in completeness studies.

5.1.4.2.4 Setup 4

Setup 4, with training samples in $[prop, 100\%]$ and test samples in $[100\%]$, has its graphs shown for Pixel-Based Baseline Classification, in Figure B.7 and Figure B.8. Also, an illustration of this Setup classes' pattern is presented in Figure A.4.

These results, Similarly to Setup 2, show that all classes tend to converge to accuracy of 100% in SVM-OAO. However, KNN-5 seems to be more sensitive to the quality of training samples, mainly concerning Forest and Crops. Forest KNN-5 UA varied from 37.18% to 89.18% while Crops varied from 57.86% to 92.22%, indicating substantial confusion between these two classes for lower pixel qualities, which is related to KNN-5 ability to divide the feature space according to the classes. An interesting point is that even though training classes was considered pure, UA did not approximate to 100% as other classes did, emphasising this KNN-5 drawback.

Analysing KNN-5 PA, Water and Bare Soil converged to 100% though Water had and increasing range from 67.15% to 98.53%. When it comes to Crops, it varied from 67.82% to 90.15% and Forest varied from 43.27% to 91.29%. This also emphasises the classifier response to the quality of training data.

5.1.4.2.5 Setup 5

Setup 5, with training samples in $[prop, prop + 5\%[$ and test samples in $[prop, 100\%]$, has its graphs shown for Pixel-Based Baseline Classification, in Figure B.9 and Figure B.10. Also, an illustration of this Setup classes' pattern is presented in Figure A.5. Unlike other Setups, Setup 5 presented somewhat similar results for both classifiers, indicating that the growth of quality of training and test samples infer in a good completeness.

All classes converged their accuracies to 100%. Regarding user accuracy, Forest, Bare Soil and Water showed consistent class reliability for all *prop* values in both classifiers. Crops, on the other hand, grew from 44.95% to 99.94% (KNN-5 UA) and from 13.72% to 99.94% (SVM-OAO UA), showing that SVM-OAO seems to be more sensitive to training purity when there is mixed testing samples in this specific situation.

When it comes to producer accuracy, Water presented its values close to 100%. Forest increased from 76.05% to 100.00% (KNN-5 PA) and 68.42% to 100.00% (SVM-OAO PA), proving the interference of non-homogeneous classes in this class' accuracy. Bare Soil also grew from 76.03% to 100.00% (KNN-5 PA) and from 70.28% to 100.00% (SVM-OAO PA), highlighting the conclusion regarding class homogeneity. Finally, Crops had its variation in between 68.31% and 99.72% (KNN-5 PA) and 96.56% and 99.83% (SVM-OAO PA).

All results indicate the interference of reference data quality on spatial data quality, mainly for lower *prop*.

5.1.4.2.6 Setup 6

Setup 6, with training samples in $[prop, 100\%]$ and test samples in $[prop, 100\%]$, has its graphs shown for Pixel-Based Baseline Classification, in Figure B.11 and Figure B.12. Also, an illustration of this Setup classes' pattern is presented in Figure A.6.

In general, all classes presented increasing accuracies converging to 100%, with minimum accuracy value of $\approx 68\%$ (Crops). This Setup indicates that, regardless of the class heterogeneity, the higher the quality of training and test sample, the higher the accuracy.

5.2 Region-based baseline classification

5.2.1 Baseline classification

The second baseline classification used region-based segmentation followed by Decision Trees 5.0 classifier using *eCognition* 9.1 and TerraView (Trimble Germany GmbH, 2014; INPE, 2020). For the segmentation, the input parameters were $scale = 50$, $compactness = 0.5$ and $shape = 0.3$. Also, all possible object (and all possible angles) features were extracted, being: area, asymmetry, average branch length, average length of edges, average area represented by segments, border index, border length, brightness, compactness, curvature, degree of skeleton branching, density, distance to scene border, elliptic fit, GLCM ang. 2^{nd} moment, GLCM contrast, GLCM correlation, GLCM dissimilarity, GLCM entropy, GLCM homogeneity, GLCM mean, GLCM standard deviation, GLDV ang. 2^{nd} moment, GLDV contrast, GLDV entropy, GLDV mean, is 3D, is at active pixel, is connected, length, length of longest edge, length of main line, thickness, width, level, level number, main direction, max. diff., maximum branch length, mean layer (1, 2, 3, 4), number of edges, number of higher levels, number of neighbours, number of pixels, number of segments, number of sublevels, number of sub-objects, perimeter, polygon sel-intersection, radius of largest/smallest enclosing, rectangular fit, rel. border to image border, roundness, shape index, standard deviation (layer 1, 2, 3, and 4), standard deviation curvature, standard deviation of area represented by segment, standard deviation of length of edges.

Once the image was segmented and had the features extracted, the region samples for training were selected manually, considering stratified sampling with 160 samples per class as shown in Table 5.9, in a total of 640 training samples. Regarding test samples, they were randomly selected as 30% of total Pixel-based samples and this is the reason why there are more test samples than training samples, in a total of $28,667$ samples. The training regions are shown in Figure 5.8. We point out that the same analyst selected both sets using the same criteria.

Table 5.9 - Region-based Baseline Classification number of training samples per class.

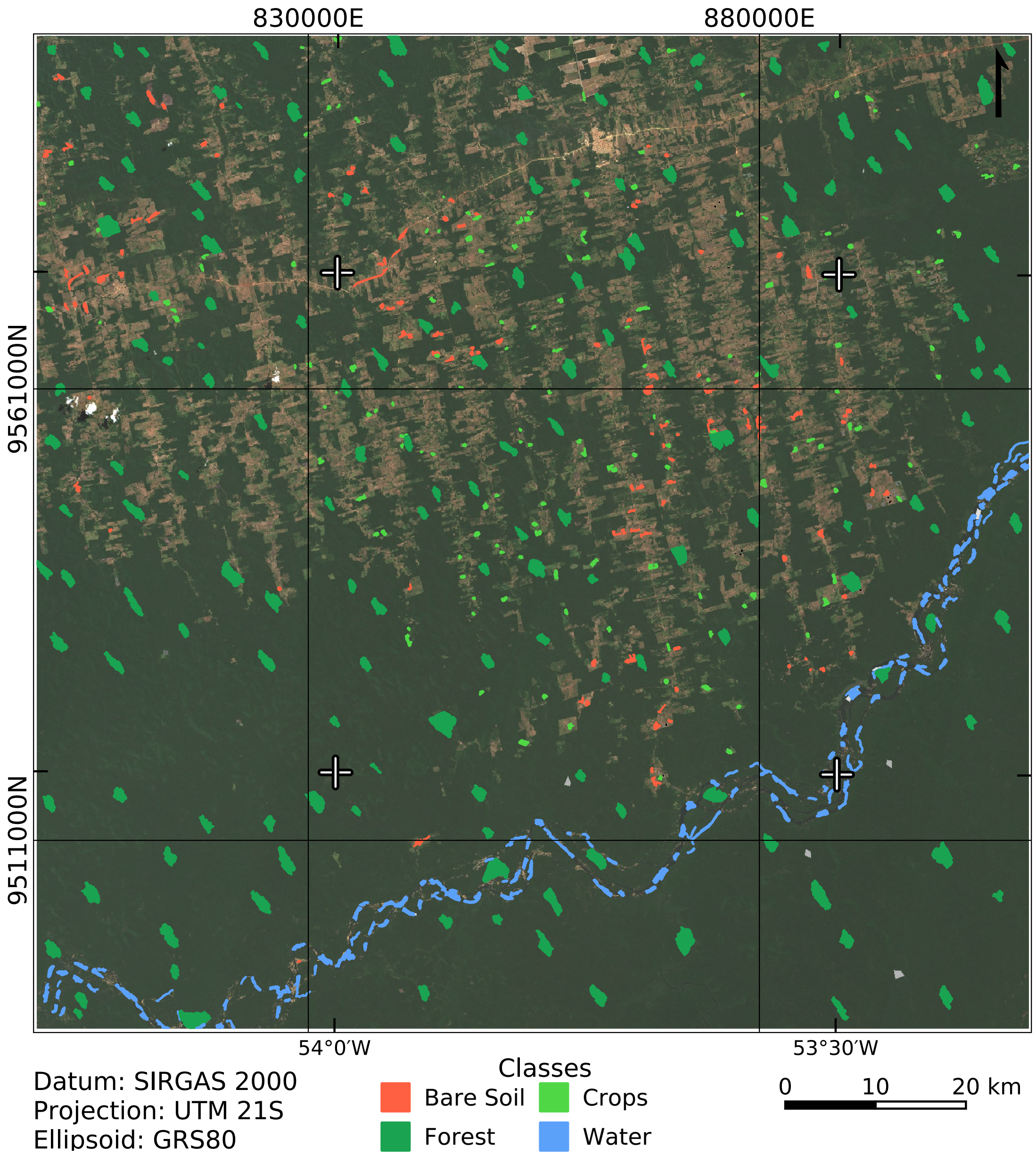
Class	Forest	Crops	Bare Soil	Water	Total
Number of training samples	160	160	160	160	640
Number of test samples	6,068	8,529	6,097	7,973	28,667

SOURCE: Author.

The training samples followed by the image classification led to the confusion matrix presented in Table 5.10. From the confusion matrix, the computed thematic accuracy showed a *kappa* index of 0.82619 with an OA of 86.933%, hence Condition IV is satisfied. This thematic accuracy is lower than the one from Pixel-based Baseline image classification (PIX) (Section 5.1.1) even though when visually analysing the resulting classified map, this classification seems more coherent with the original image, as shown in Figure 5.9. We point out that the noise presented in this classification is substantially smaller than the one from the Pixel-based Baseline classification.

A point to be addressed regarding the thematic accuracy is that the object for training the classifier is different than the object for testing it. Therefore, this analysis is purely cartographic and not over the classifier. Assuming that image segmentation followed by classification generally presents better outcomes than PIX we can suppose that the training data quality was better than the test data quality. In this sense, a likely reason for this thematic accuracy being smaller than the one from Pixel-based Baseline Classification is that the quality of the test data influences more directly the thematic accuracy than training data, as previously discussed in Section 5.1.4. Another probability is that when we segment an image in several region-objects, border pixels will necessarily occur in the border of segments, which leads to a less noisy classification, hence more homogeneous pixels.

Figure 5.8 - Region-based Baseline Classification using Decision Trees - training objects.



Training objects for region-based presented over a Sentinel-2A TCI image from 09 August 2020. This image with ROI in Black and White (BW) is shown in Figure C.4, page 200. Crops and Bare soil had their objects resized so they could be seem from the used scale.

SOURCE: Author.

Table 5.10 - Region-based Baseline Classification using Decision Trees - Confusion Matrix.

		Predicted				
	Class	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Truth	**Forest**	**6068**	0	0	0	6068
	Crops	966	**5354**	2209	0	8529
	Bare Soil	0	571	**5526**	0	6097
	Water	0	0	0	**7973**	7973
	Total	7034	5925	7735	7973	**28667**

SOURCE: Author.

Regarding the completeness elements, user and producer accuracies are shown in Table 5.11. When it comes to user accuracy (UA), i.e. when samples from another class are mistakenly assigned to that class, Forest and Water presented no errors while Crops presented $UA = 62.774\%$ and Bare Soil, 91.635%. For producer accuracy (PA), Water samples were not omitted (100.00% PA), Crops showed 90.363% and Forest, 86.627%; Bare Soil had the lowest PA value of 71.441%. Comparing the confusion, Crops were confused with Forest and Bare Soil though in a higher rate than PIX. Besides the likely reason points out, there is also the possibility of the segmentation using different criteria for determining border pixels between these three classes which may have influenced on these outcomes.

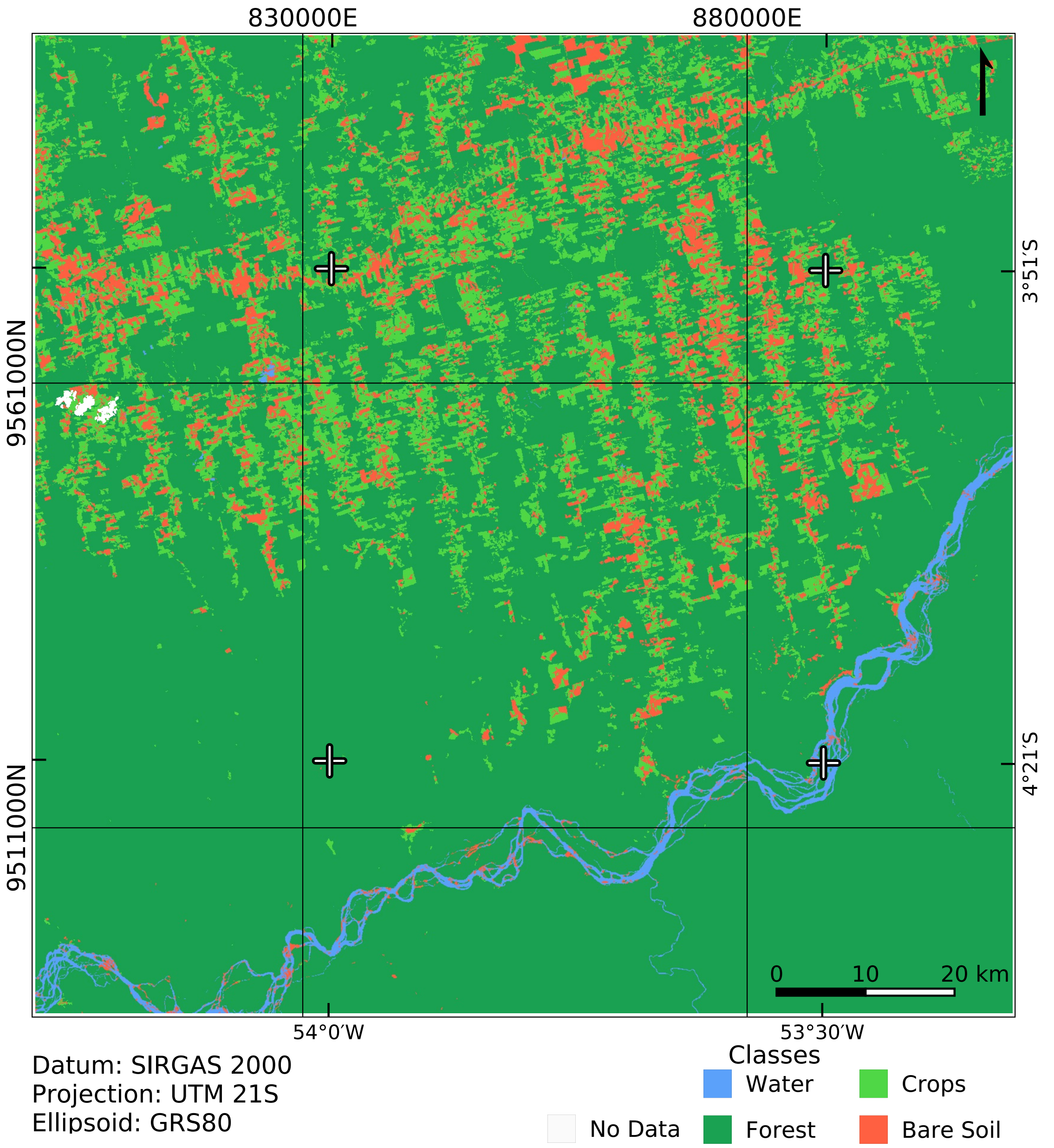
At this point, Conditions, I, II, III and IV have all been met, then, it is possible to carry on to the Reference Sample Selection (RSS) process for the region-based baseline classification (REG).

Table 5.11 - Region-based Baseline Classification using Decision Trees - Completeness.

Class	**User accuracy (%)**	**Producer accuracy (%)**
Forest	100.000	86.267
Crops	62.774	90.363
Bare Soil	90.635	71.441
Water	100.000	100.000

SOURCE: Author.

Figure 5.9 - Region based First Classification using Decision Trees - classified image.



White pixels correspond to NoData values. This image in Black and White is on Figure C.5, page 201.

SOURCE: Author.

5.2.2 Reference Sample Selection part I - filtering candidate samples

Similarly to PIX, the data is presented in four formats: histogram, pie charts, table and frequency map, each one providing a distinct interpretation of the data.

The first one is the histogram (Figure 5.10), divided into histogram of accumulated modal class proportion, assuming the set $[prop, 100\%]$ and histogram of samples per stratum in the set $[prop, prop + 5\%[$. These histograms are in logarithmic scale due to the amount of Forest class data in $prop = 100\%$. Once again, Forest had the great majority of its candidate samples in pure quality. Meanwhile, Crops, Bare Soil and Water presented a more flattened pattern, showing a better distribution of their candidate data over all *prop* values. In this sense, these graphs reinforce that the class representativity varies according to sample quality.

The second are the pie charts, in Figure 5.11 visually showing the variability of the class' representativiness as *prop* varies. These charts imply that, for REG, over 50% of the data for all classes are in $prop = 100\%$ indicating that REG may have generalised quality information.

The third format is the table (Table 5.13) with the exact amount of data and their respective percentage regarding the total number of candidate samples per class. Firstly analysing Forest class, 94.43% of this class is at $prop = 100\%$; while the other strata of this class present percentages under 1.00%. When it comes to the other classes, Water, Crops and Bare Soils had their percentages with pure samples in 69.04%, 65.08% and 64.34%, respectively. In spite of these three classes having a considerable more flattened distribution, the majority (over 50%) of their candidate data lays in pure pixels.

Also, from Table 5.13, we can analyse central tendency with grouped mean and *std* values, presented in Table 5.12. From this table, we see that the average values lay in $prop \geq 91\%$, showing the central mass around purer samples and Crops, Bare Soil and Water present similar *std* while Forest showed a lesser variation. When analysing the distribution, all classes present negative skew, which means that the central mass is on the right part of the chart, hence it is in accordance with Figure 5.10. Also, Forest presents a higher absolute skew value, indicating that the central mass is more to the right. Regarding kurtosis, Water presented a negative excess kurtosis, being platykurtic meaning that its peak is more in a shape of an inverted "U", hence having more distributed quality values. Crops and Bare Soil present slightly positive excess kurtosis, indicating leptokurtic distribution, meaning that theses classes have a slightly more accentuated peak, indicating less varied distribution. Finally, Forest presented substantially positive excess kurtosis, almost in a shape of an inverted "V", hence being the less distributed value. To conclude, for this baseline classification, all classes are best represented by samples with higher quality.

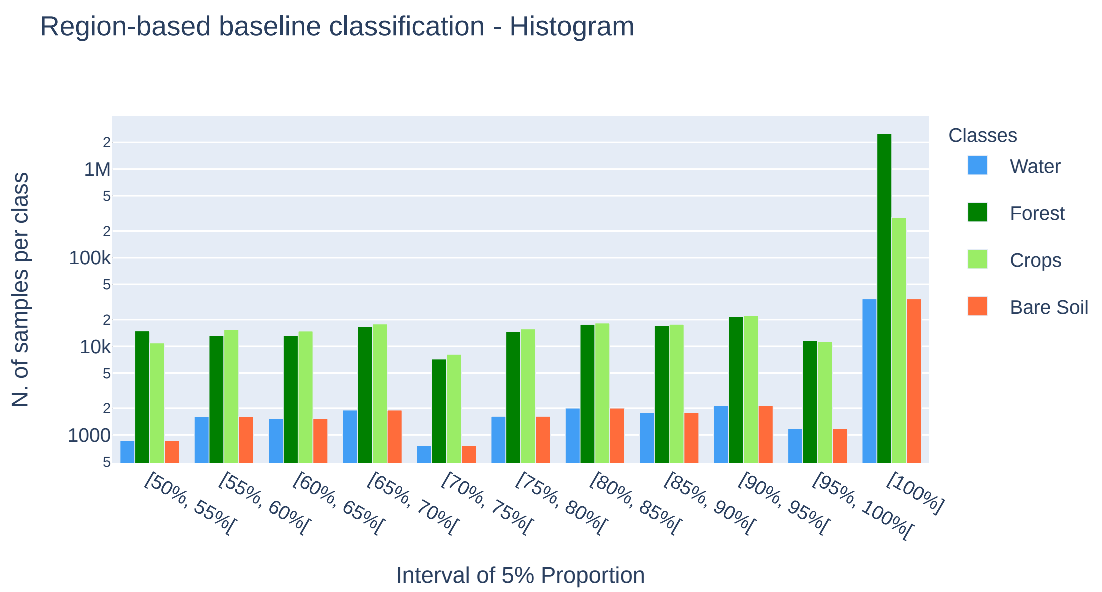
Table 5.12 - Region-based Baseline Classification using Decision Trees - Histogram Distribution Statistics.

Class	**Grouped mean** (%)	**Grouped *std*** (%)	**Skewness**	**Excess kurtosis**
Water	92.71%	13.30%	-1.71	-2.46
Forest	98.66%	6.47%	-5.40	29.63
Crops	91.69%	14.03%	-1.53	0.96
Bare Soil	91.51%	14.10%	-1.490	0.81

SOURCE: Author.

When comparing these results with the ones from PIX (Table 5.6), there is a substantial difference on percentage of Bare Soil (27.32% to 65.08%) and Crops (21.81% to 64.34%) classes. This fact, is likely related to inherent characteristics of image segmentation followed by classification. Additionally, this fact may affect how we understand classes representativity, once it seems to vary according to the baseline image. Apparently, the definition of class representativiness depends on what criteria we use to define that certain class for that certain area, instead of having a defined pattern for classifications, as shown in Table 5.5 and Table 5.12. For example, when the analyst selects points/pixels as classification object, they may need to consider a broader variety of quality; however in the case of segmentation, the analyst may want to select purer objects.

Figure 5.10 - Region-based based First Classification using Decision Trees - histogram.



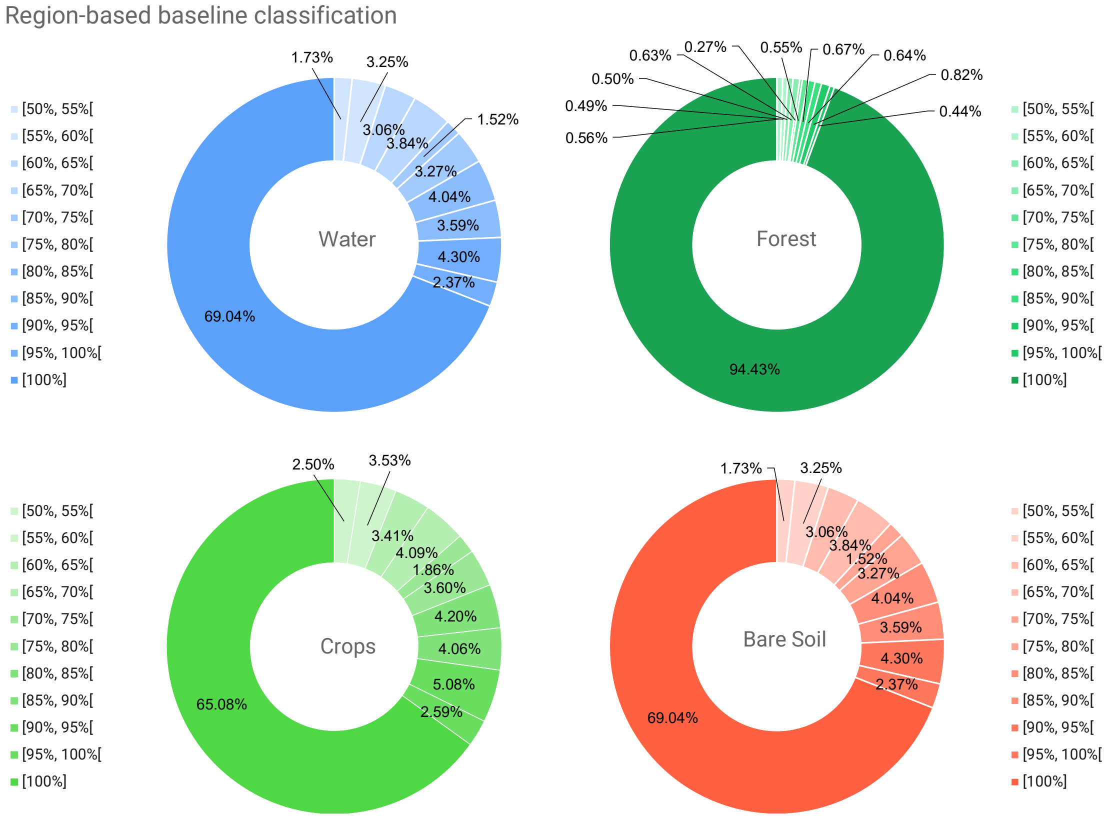
Number of pixels presented on the superior image in accumulated samples per frequency on set $[prop, 100\%]$ and in the inferior image with samples per interval of frequency $[prop, prop + 5\%[$.

SOURCE: Author.

Finally, the third format shows the spatial distribution of the candidate samples in Figure 5.12, with summarised strata for a better image visualisation, in the set $[prop, prop + 10\%[\ \forall\ prop \in \{50\%, 60\%, 70\%, 80\%, 90\%, 100\%\}$ with the increasing values from lighter to darker colours. This format is generalised for all classes, showing explicitly the modal class proportion and not the land cover class defined as most frequent class. For the REG image, border pixels seem visually more delineated. Additionally, it is possible to observe more clearly the fishbones, which are areas more likely to present Crops and Bare Soil classes. Even so, image border pixels in this situation are related to border from the segments, hence this classification tends to have more homogeneous areas.

Therefore, this baseline classification generated different results from the Pixel-based Baseline Classification, specially when it comes to number of candidate samples per stratum. A plausible reason for that lies in the segmentation characteristics that generate a less noisy classified image. These characteristics may also affect the study of class representativity, which may influence further studies of the RSS process.

Figure 5.11 - Region-based Baseline Classification using Random Forest - pie charts of distributed data per class.



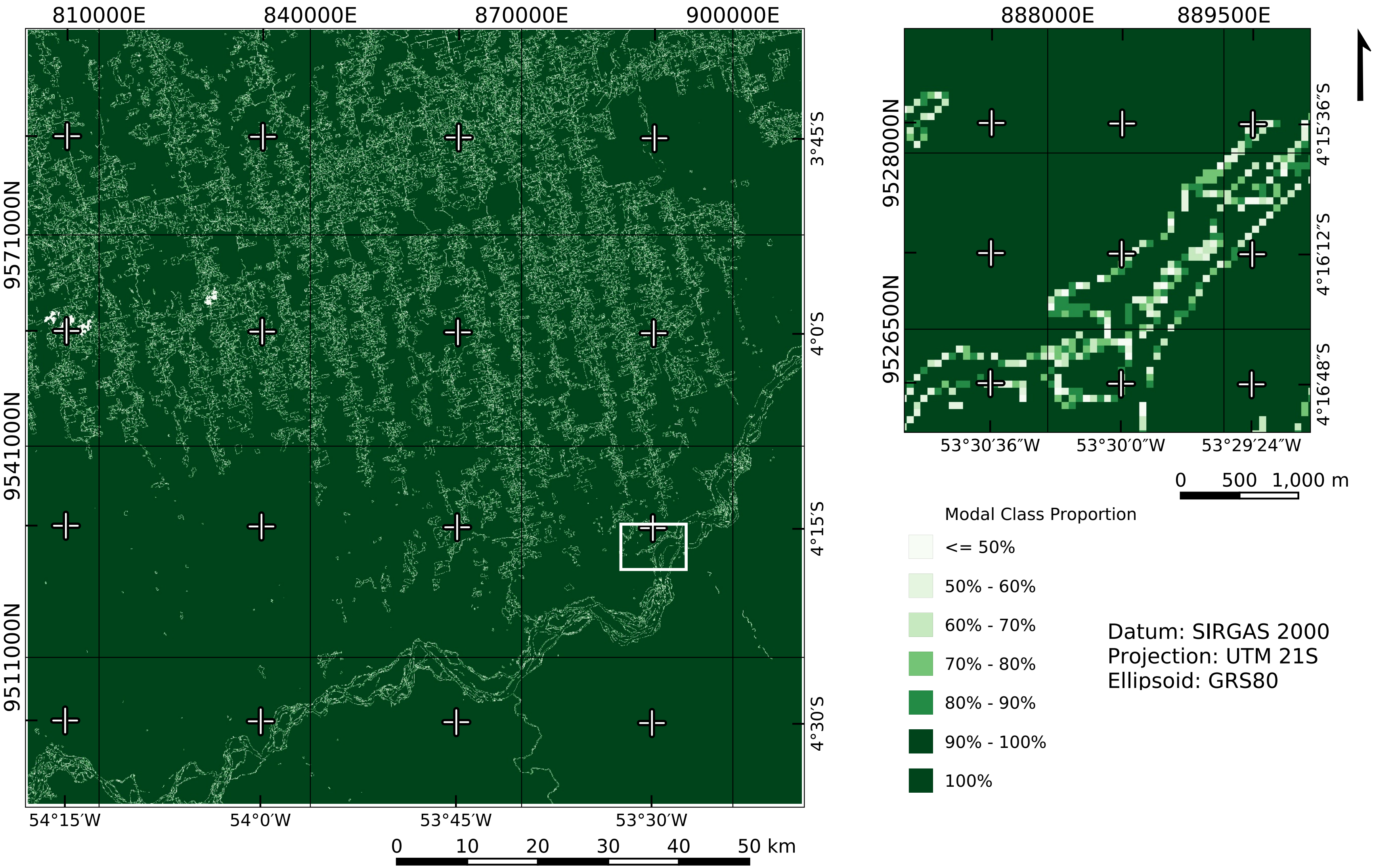
SOURCE: Author.

Table 5.13 - Region-based based First Classification - number of samples per stratum and their respective percentage per class.

prop		Number of samples per stratum									
		Water		Forest		Crops		Bare Soil		**Total**	
		N.	(%)	N.	(%)	N.	(%)	N.	(%)	N.	(%)
50%	55%	858	1.73%	14,916	0.56%	10,905	2.50%	4,639	2.31%	31,318	0.94%
55%	60%	1,614	3.25%	13,141	0.49%	15,386	3.53%	7,532	3.75%	37,673	1.13%
60%	65%	1,518	3.06%	13,206	0.50%	14,867	3.41%	7,298	3.63%	36,889	1.10%
65%	70%	1,907	3.84%	16,655	0.63%	17,863	4.09%	8,403	4.18%	44,828	1.34%
70%	75%	756	1.52%	7,184	0.27%	8,129	1.86%	3,835	1.91%	19,904	0.60%
75%	80%	1,622	3.27%	14,721	0.55%	15,704	3.60%	7,453	3.71%	39,500	1.18%
80%	85%	2,006	4.04%	17,697	0.67%	18,309	4.20%	8,713	4.34%	46,725	1.40%
85%	90%	1,781	3.59%	17,012	0.64%	17,725	4.06%	8,265	4.11%	44,783	1.34%
90%	95%	2,133	4.30%	21,687	0.82%	22,164	5.08%	10,323	5.14%	56,307	1.68%
95%	100%	1,179	2.37%	11,591	0.44%	11,290	2.59%	5,199	2.59%	29,259	0.88%
100%		34,276	69.04%	2,507,228	94.43%	283,892	65.08%	129,292	64.34%	2,954,688	88.41%
Total		**49,650**	100.00%	**2,655,038**	100.00%	**436,234**	100.00%	**200,952**	100.00%	**3,341,874**	100.00%

SOURCE: Author.

Figure 5.12 - Region-based based Baseline Classification using Decision Trees: frequency map.



SOURCE: Author.

5.2.3 Reference Sample Selection part II - selecting reference data and image classification

RSS Part II process filters the candidate samples to form the *bag* so the reference samples can be selected, as showed in the methodology flowchart in Figure 4.2.

To determine the *bag* size, Table 5.13 is used where the lowest value in a stratum is 756 in the set $[70\%, 75\%[$. Then, the chosen value for the *bag* size was $n_{samples} = 750$ as this is the nearest round number. Consequently, there are 750 samples per stratum in the *bag*.

Furthermore, as n_{bag} in this process is the same as the one set for the PIX approach, the available reference data varies from $3,000$ $(750 \times 4_{classes} \times 1_{interval})$ to $33,000$ $(750 \times 11_{intervals} \times 4_{classes})$ samples, depending on in what *prop* the simulation is working on.

Analogously, splitting the data follows the proportion of $2/3$ and $1/3$ for training and test samples, respectively. Thus, for the four classes, $n_{train} = 500 \times 4 = 2,000$ and $n_{test} = 250 \times 4 = 1,000$.

Defined the number of training and test samples as well as n_{bag}, the image classification with the Monte Carlo simulation as sequence of the RSS - Part II approach follows.

5.2.4 Spatial data quality

Following the process presented in the methodology flowchart (Figure 4.2), the image classification using the Monte Carlo simulation was conducted, where there was 100 repetitions per modal class proportion (either set $[prop, 100\%]$, or set $[prop, prop + 5\%[$). Therefore, the total number of repetitions per Setup was $1,100$ for each classifier (KNN-5 and SVM-OAO). The spatial data quality (SDQ) considering thematic accuracy and completeness with their respective standard deviations were evaluated.

The first studied SDQ component was thematic accuracy, in Section 5.2.4.1, followed by completeness, in Section 5.2.4.2. Besides, due to the substantial number of resulting data, they are presented hereafter in Appendix B.2: firstly the charts with *kappa* index and producer and user accuracies (Appendix B.2.1), followed by the tables with their exact values and standard deviation (Appendix B.2.2) and, finally, the confusion matrices (Appendix B.2.3).

5.2.4.1 Thematic accuracy

The thematic accuracy of the Region-based Baseline Classification (REG) approach are shown in Figure 5.13 and Figure 5.14 for each Setup in the upper chart for KNN-5 classifier and in the bottom chart for SVM-OAO classifier. Despite the charts presenting the *kappa* index, the tables containing overall accuracy are in Appendix B.2.2. The charts also show errors bars. The results are also presented in tabular results in Table 5.14 and Table 5.15.

Similarly to Pixel-based Baseline Classification results, the Setups are divided into two groups: the first group has Setups 3, 4 and 6 while the second group has Setups 1, 2 and 5 concerning three mains points: (i) response of the classifier to the quality of reference data; (ii) how convergence of reference data quality affects the thematic accuracy and (iii) how training and test samples affect the thematic accuracy.

Firstly analysing thematic accuracy of the first group (Figure 5.13), with their training samples in the set $[prop, 100\%]$, all its Setups presented different patterns, varying from *kappa* ≈ 0.65 to 0.95. Finally, for Setup 6 with test in set $[prop, 100\%]$, we can perceive almost linear increase in both classifiers, increasing up to 0.90.

Setup 3, with test set in $[50\%, 100\%]$, expresses mainly how each classifier respond to increasingly divergence of training and test data qualities. An interesting fact is that *kappa* values for this Setup were all below 0.70 indicating that, when using mixed quality as test samples, we may not achieve exceptional results. As this divergence increased, so did the thematic accuracy, showing that the thematic accuracy may not be so dependent on training samples quality even though their quality increased.

Studying Setup 4, with test in $[100\%]$, the first point to observe is that its thematic accuracy was higher than the one from Setup 3. The differences between these Setups are in the test data quality and its lowest *kappa* value was ≈ 0.85. This fact leads us to the understanding that test sample quality in fact affect the thematic accuracy more dramatically. Still, for this baseline experiment, training samples seemed to have some influence on the thematic accuracy, once *kappa* varied from 0.85 to 0.90.

Finally, for Setup 6 with test in set $[prop, 100\%]$, we can perceive an almost linear increase in both classifiers, increasing up to 0.90. There are two likely analyses regarding this Setup: the first is that the thematic accuracy increased alongside with the analyst experience, hence the higher the analyst's experience, the better. The other point of view is related to the quality of test samples that increased, leading to a higher thematic accuracy.

Comparing these three Setups from the first group, a funny fact is that Setup 6 starts together with Setup 3 and finished together with Setup 4. Clearly, the reason is that the qualities of training and test samples for these two points is the same. Another meaningful point of comparison is the difference in the thematic accuracy for same *prop* values. The difference in treatment was exclusively in the quality of training data and yet, both classifiers presented the similar response pattern. Therefore, answering the three analysis points: (i) there was no significant difference in the classifier response to these three Setups; (ii) for this baseline classification, the convergence of the quality between training and test data has some influence though it is not so symbolic and (iii) the thematic accuracy is clearly very sensible to the test samples quality.

Moving on to the second group (Figure 5.14), all of them presented training set in $[prop, prop + 5\%[$, hence they are not realistic. Analysing firstly Setup 1, with test samples in $[50\%, 100\%]$, it showed the lowest values for this group and also the lowest *kappa* range. This Setup explains mostly convergence of the qualities of training and test data: the curve is close to an inverted "U", indicating that for when they diverge the most, the thematic accuracy is lower and when they converge the accuracy is higher. The convergence, in this Setup, relates to training samples in their average, *prop* around 70% and 75%, which is around the average of quality between 50% and 100%. Besides that, this Setup also indicates how each classifier responds to this divergence: KNN-5 showed a higher *kappa* variation which may be related to the ability to separate the classes in the feature space.

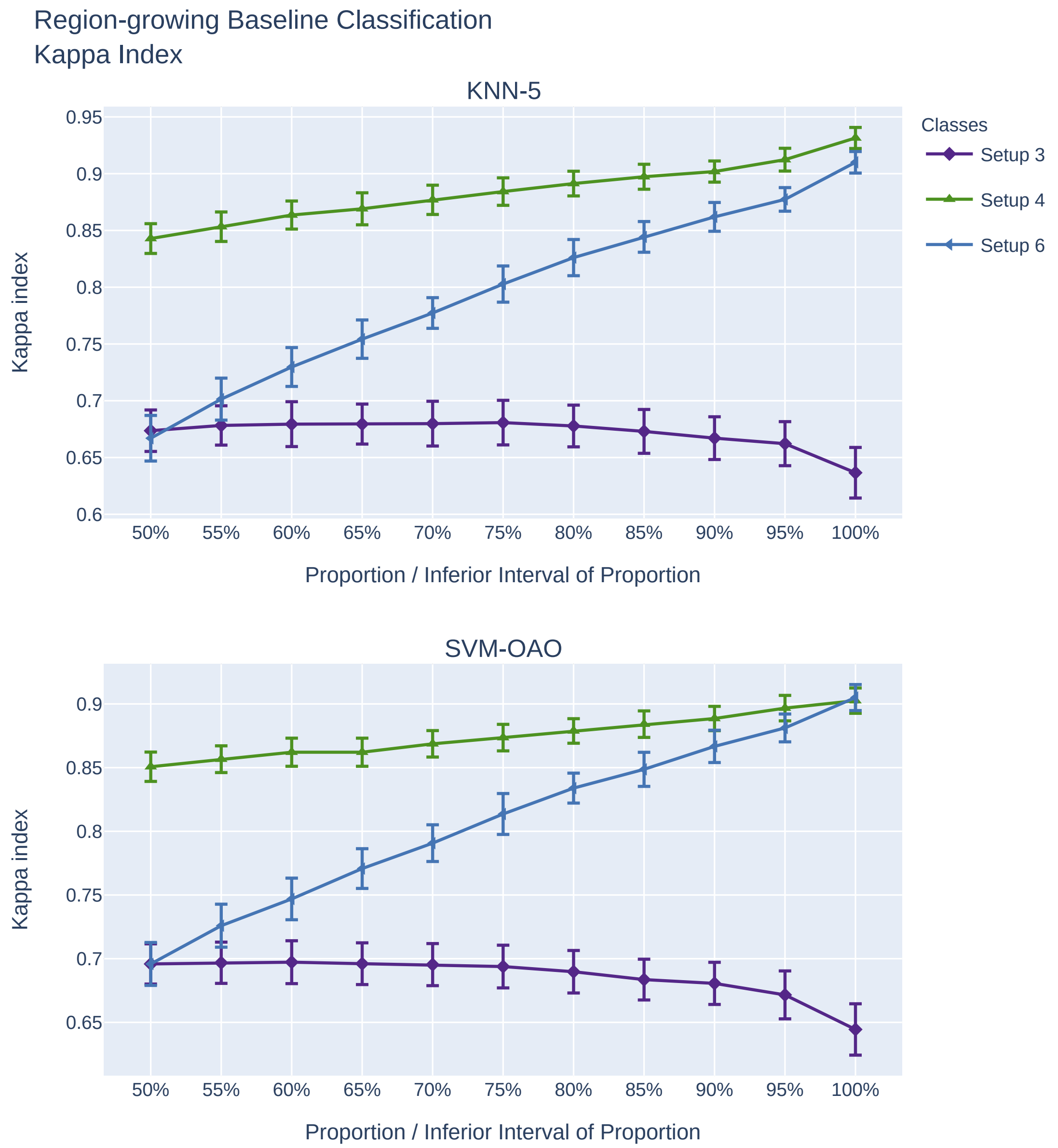
Setup 2 with test set in [100%] presented the highest *kappa* values for this group although it showed a growth according to the quality of training samples. Also, differently from this Setup in Pixel-based Baseline Classification, there was an actual variation of *kappa* index. This indicates that the defining criteria of pixel quality in this experiment infer more influent training samples for the thematic accuracy. Nevertheless, the quality of test samples still has an impact on that. If we compare the classifiers, we can see that SVM-OAO presented, in general, higher *kappa*, mainly for lower *prop* values, expressing that it may deal better with separating the feature space than KNN-5.

The last one is Setup 5 with its test in the set $[prop, 100\%]$. We can see a linear increase of the thematic accuracy as the quality of training samples increases and that there is a more homogeneous quality of test samples. Regarding this Setup, we cannot take any specific conclusions in spite of its importance to general assessment.

To compare these three Setups, again, we see clearly the effect of test samples on the thematic accuracy. Setups 1 and 2 present a fixed quality of test data while Setup 3 varies it. If we observe the difference between Setups 1 and 2, we see the discrepancies of *kappa* values and this verifies the sensibility of the thematic accuracy regarding test samples. Another point is that SVM-OAO generally presented more accurate results than KNN-5 even though both classifiers presented similar curve patterns for the three Setups. The third point is that both classifiers responded poorly to divergence of training and test data in spite of the used *prop*.

To study the effect of training data, we can also compare these pair of Setups: 1 and 3, 2 and 4, 5 and 6. All the pairs tended to converge their *kappa* as *prop* increased and there was higher difference in the thematic accuracy for $prop \leq 75\%$. This corroborates with the idea that test samples tend to have stronger influence on the thematic accuracy than training samples. However, the lower the sample quality, the higher the influence of training samples on thematic accuracy. Additionally, when comparing the classifiers, SVM-OAO presented higher accuracy values, but both classifiers presented similar curve patterns for all Setups in this baseline classification. This means that the classifier has some influence on the thematic accuracy though the variation of sample quality seems to be stronger.

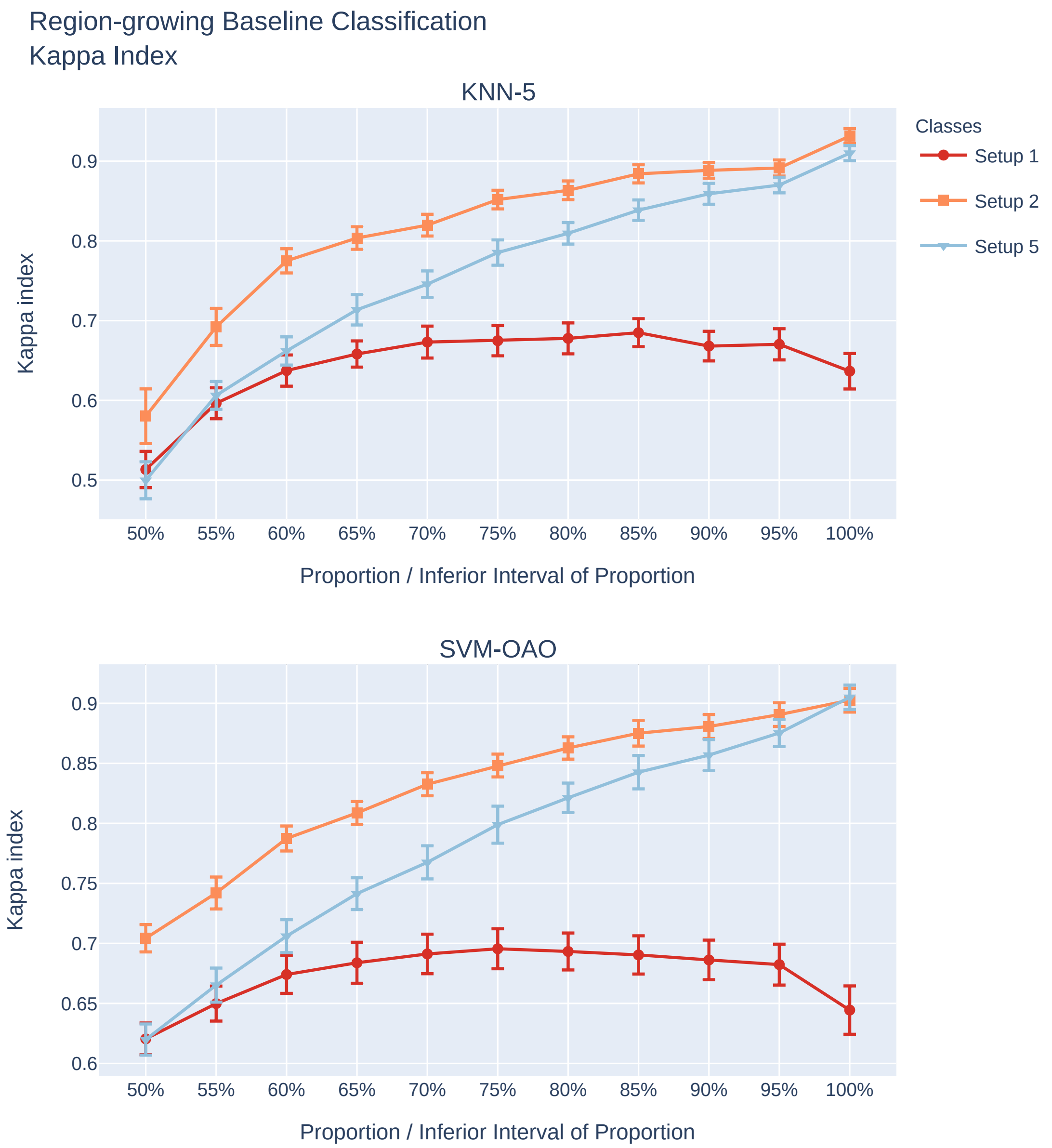
Figure 5.13 - Region-based Baseline Classification using Decision Trees C5.0 - Thematic Accuracy with error bars for Setups 3, 4 and 6.



Where Setup 3 corresponds to train: $[prop, 100\%]$ / test: $[50\%, 100\%]$ and Setup 4 corresponds to train: $[prop, 100\%]$ / test: $[100\%]$ and Setup 6 corresponds to train: $[prop, 100\%]$ / test: $[prop, 100\%]$. More detailed results are presented in Section B.2.

SOURCE: Author.

Figure 5.14 - Region-based Baseline Classification using Decision Trees C5.0 - Thematic Accuracy with error bars for Setups 1, 2 and 5.



Where Setup 1 corresponds to train: $[prop, prop + 5\%]$ / test: $[50\%, 100\%]$; Setup 2 corresponds to train: $[prop, prop + 5\%]$ / test: $[100\%]$ and Setup 5 corresponds to train: $[prop, prop + 5\%[$ / test: $[prop, 100\%]$. More detailed results are presented in Section B.2.

SOURCE: Author.

Table 5.14 - Region-based Baseline Classification - Thematic accuracy and standard deviation for KNN-5.

prop	**Setup 1**		**Setup 2**		**Setup 3**		**Setup 4**		**Setup 5**		**Setup 6**	
	kappa	std	*kappa*	std	*kappa*	std	*kappa*	std	*kappa*	std	*kappa*	std
50%	0.513	0.023	0.580	0.034	0.674	0.018	0.843	0.013	0.500	0.023	0.667	0.020
55%	0.596	0.019	0.692	0.023	0.678	0.017	0.853	0.013	0.606	0.017	0.701	0.018
60%	0.637	0.020	0.775	0.015	0.679	0.020	0.864	0.012	0.662	0.018	0.730	0.017
65%	0.658	0.016	0.804	0.014	0.679	0.018	0.869	0.014	0.714	0.019	0.754	0.017
70%	0.673	0.020	0.820	0.014	0.680	0.020	0.877	0.013	0.746	0.017	0.777	0.014
75%	0.675	0.019	0.852	0.012	0.681	0.020	0.884	0.012	0.785	0.016	0.803	0.016
80%	0.678	0.019	0.863	0.012	0.678	0.018	0.891	0.011	0.809	0.013	0.826	0.016
85%	0.685	0.018	0.884	0.011	0.673	0.019	0.897	0.011	0.839	0.013	0.844	0.014
90%	0.668	0.019	0.888	0.010	0.667	0.019	0.902	0.009	0.859	0.013	0.862	0.013
95%	0.670	0.020	0.891	0.010	0.662	0.019	0.912	0.010	0.870	0.010	0.877	0.010
100%	0.637	0.022	0.931	0.009	0.637	0.022	0.931	0.009	0.910	0.009	0.910	0.009

SOURCE: Author.

Table 5.15 - Region-based Baseline Classification - Thematic accuracy and standard deviation for SVM-OAO.

prop	**Setup 1**		**Setup 2**		**Setup 3**		**Setup 4**		**Setup 5**		**Setup 6**	
	kappa	std	*kappa*	std	*kappa*	std	*kappa*	std	*kappa*	std	*kappa*	std
50%	0.620	0.013	0.704	0.011	0.696	0.016	0.851	0.012	0.620	0.013	0.696	0.017
55%	0.650	0.015	0.742	0.013	0.697	0.016	0.857	0.010	0.665	0.014	0.726	0.017
60%	0.674	0.016	0.787	0.010	0.697	0.017	0.862	0.011	0.706	0.014	0.747	0.016
65%	0.684	0.017	0.809	0.009	0.696	0.016	0.862	0.011	0.741	0.013	0.771	0.016
70%	0.691	0.016	0.833	0.010	0.695	0.017	0.869	0.010	0.768	0.014	0.791	0.014
75%	0.696	0.017	0.848	0.010	0.694	0.017	0.874	0.010	0.799	0.015	0.814	0.016
80%	0.693	0.015	0.863	0.009	0.690	0.017	0.879	0.010	0.821	0.012	0.834	0.012
85%	0.690	0.016	0.875	0.011	0.684	0.016	0.884	0.010	0.843	0.014	0.849	0.013
90%	0.686	0.016	0.881	0.010	0.681	0.017	0.889	0.010	0.857	0.013	0.867	0.013
95%	0.682	0.017	0.891	0.010	0.672	0.019	0.897	0.010	0.875	0.011	0.881	0.011
100%	0.644	0.020	0.903	0.010	0.644	0.020	0.903	0.010	0.905	0.010	0.905	0.010

SOURCE: Author.

5.2.4.2 Completeness

The completeness elements will study thematic accuracy results of the RSS process for REG, by analysing results for each class using user accuracy and producer accuracy. Due to the amount of resulting charts (12), they are presented in Appendix B.2.1, the tables are in Appendix B.2.2 and confusion matrices in Appendix B.2.3 alongside with the thematic accuracy per Setup.

5.2.4.2.1 Setup 1

This Setup has training samples in $[prop, prop + 5\%[$ and its test samples are in $[50\%, 100\%]$. Its resulting charts are in Figure B.13 and Figure B.14. Additionally, an illustration of Setup 1 feature space is presented in Figure A.1.

Firstly, concerning Forest class, both classifiers responded differently for UA and PA. Regarding user accuracy, Forest varied from 70.67%, rose up to 82.36% ($prop = 70\%$) then fell to 53.16% in KNN-5 while, for SVM-OAO it fell from 91.37$ to 60.13% showing that each classifier responds differently to this class when there is a variation of reference data. When it comes to producer accuracy, it grew from 58.39% to 88.35% (KNN-5 PA) and from 61.93% to 86.19% (SVM-OAO PA). Similarly to Water, this class is understood to be spectrally homogeneous and lower PA indicated some heterogeneity, which is uncommon.

For Crops class, its UA rose for both classifiers, varying from 34.09% to 79.52$ (KNN-5 UA) and from 10.86% to 76.38% (SVM-OAO UA), showing the lack of reliability of this class for lower *prop* values. When it comes to producer accuracy, Crops varied from 39.16%, rose up to 61.52% ($prop = 70\%$) then fell to 51.77% in the case of KNN-5 PA. For SVM-OAO PA, it fell from 77.69% to 53.20%, showing that other classes interfere with this class' PA, which happened for PIX studies. These outcomes also show the difference of how the classifiers respond to this class considering variety in quality of reference data.

When it comes to Bare Soil, firstly analysing user accuracy, it varied from 54.31% to 75.84% ($prop = 65\%$) then fell to 65.41% in KNN-5. Now, SVM-OAO UA varied from 85.10% to 66.10%. Together with Crops class, we may suggest that there was confusion between these two classes and, as segmentation tends to generalise the pixel quality information, this may have affected this situation. Moving on to producer accuracy, KNN-5 PA grew from 64.64% to 73.77% and, for SVM-OAO PA, rose from 66.68% to 73.61% ($prop = 80\%$) then fell to 72.19%. In general, Bare

Soil varied its completeness around 70%, which may be related to the definition of sample quality, that segmentation infers to worse values.

For Water, it presented the highest completeness values. Concerning user accuracy, it varied from 94.94% to 92.89% (KNN-5 UA) and from 98.82% to 90.73% (SVM-OAO UA). Indicating that the generalisation of segments for this class was more trustworthy. Analysing samples mistakenly assigned to Water, it also showed good accuracy, presenting values of PA from 87.99% to 95.20% (KNN-5 PA) and from 89.14% to 95.40%. Therefore, this class seemed to have discrepant generalised value, when compared to other classes.

5.2.4.2.2 Setup 2

This Setup has training samples in $[prop, prop + 5\%[$ and its test samples are in $[100\%]$. Its resulting charts are in Figure B.15 and Figure B.16. Additionally, an illustration of Setup 2 feature space is presented in Figure A.2. In general, Forest and Water converged to 100.00% for user accuracy. The remaining class presented variations in their accuracy assessment values.

Bare Soil varied from 57.86% to 92.22% for KNN-5 UA and, for SVM-OA UA, it fell from 94.33% to 90.51%. These values indicate that the generalisation of regions may have increased Bare Soil homogeneity for pure test samples. Analysing producer accuracy, its variations were from 67.82% to 90.15% (KNN-5 PA) and from 69.38% to 84.80% (SVM-OAO PA).

Crops presented higher discrepancy in UA values: grew from 37.18% to 89.18% (KNN-5 UA) and from 18.07% up to 81.77% (SVM-OAO UA). Regarding producer accuracy, Crops rose in KNN-5 (43.27% to 91.29%) and in SVM-OAO, it fell from 94.82% to 89.64%., showing that SVM-OAO was able to deal better with *prop* variation.

5.2.4.2.3 Setup 3

This Setup has training samples in $[prop, 100\%]$ and its test samples are in $[50\%, 100\%]$. Its resulting charts are in Figure B.17 and Figure B.18. Additionally, an illustration of Setup 3 feature space is presented in Figure A.3. In general, user accuracy presented similar patterns for all classes for both classifiers, indicating that, for accumulated quality of training samples against mixed text quality, both classifiers respond similarly, unlike Setup 1.

Forest had its user accuracy decrease from 78.30% to 53.16% (KNN-5) and from 83.37$ to 61.13% inferring that spectral dispersion of other classes influences directly this class' reliability. When it comes to producer accuracy, this class varied from 74.73% to 88.35% (KNN-5 PA) and from 74.76% to 86.16% (SVM-OAO PA). This elucidates the fact that this class tends to be best separated in the feature space as *prop* increases, but other classes are mistakenly assigned to it as a drawback.

Crops rose from 55.44% to 79.52% (KNN-5 UA) and from 49.12% to 76.38% (SVM-OAO UA), showing that this class reliability increased as *prop* rose although it did not reach 100%. For producer accuracy, it fell from 59.97% to 51.77% (KNN-5 PA) and from 65.62% to 53.20% (SVM-OAO PA), implying that, the higher the quality of training samples, if test samples are mixed, heterogeneous classes tend to be mistakenly assigned to this class.

Bare Soil had its user accuracy varying from 72.50% to 65.41% (KNN-5 UA) and from 79.00% to 66.10% (SVM-OAO UA), showing that, as Setup 1, as *porp* increased, UA had a slight decrease for both classifiers, indicating that the classification did not separate this heterogeneous class for higher quality samples.

Water presented general high values: its user accuracy varied form 95.84% to 92.89% (KNN-5 UA) and from 97.28% to 90.73% (SVM-OAO). And for producer accuracy, it varied from 93.20% to 95.20% (KNN-5 PA) and from 92.78% to 95.40% (SVM-OAO PA). Emphasising this class' homogeneity in both aspects.

5.2.4.2.4 Setup 4

This Setup has training samples in $[prop, 100\%]$ and its test samples are in $[100\%]$. Its resulting charts are in Figure B.19 and Figure B.20. Additionally, an illustration of Setup 4 feature space is presented in Figure A.4.

Regarding Forest and Water, both classes presented all their UA near 100% for both classifiers. Nonetheless, their PA varied: Water kept accuracy close to 100% while Forest had a growing pattern: from 87.49% to 98.53% (KNN-5 PA) and from 88.16% to 97.12% (SVM-OAO PA). These values points out these classes' reliability against pure test samples, indicating their homogeneity, as previously mentioned.

Crops presented a growth for its user accuracy, varying its KNN-5 from 64.62% to 89.18% and its SVM-OAO from 60.34% to 81.77%, showing almost a linear correlation between *prop* and UA, therefore, considering pure test samples, we can rely better in this class. This reliability changes when we analyse producer accuracy; this class varied from 87.70% to 91.29% (KNN-5 PA) and from 95.93% to 89.64%. This indicates that higher *prop* values for this class affects differently both classifiers' responses.

Finally, Bare Soil did not present substantial changes in its user accuracy, with variations in between 89.51% and 92.22% (KNN-5 UA) and in between 95.33% and 90.51% (SVM-OAO UA). This shows that this class in not highly dependent on the quality training samples when testing the image classification with pure samples, indicating that both classifiers may rely more in mixed quality training samples for separating their feature space.

5.2.4.2.5 Setup 5

This Setup has training samples in $[prop, prop + 5\%[$ and its test samples are in $prop, [100\%]$. Its resulting charts are in Figure B.21 and Figure B.22. Additionally, an illustration of Setup 5 feature space is presented in Figure A.5. In this Setup, there was difference in the classifier's responses. Water kept with the highest PA and UA values with its lower value in KNN-5 PA (87.83%).

Forest presented growing or unchanging pattern for all cases. Beginning with user accuracy, it rose from 69.68% to 97.84% (KNN-5 UA) and varied from 99.68% to 99.01% (SVM-OAO UA). This showes that KNN-5 seems to be more susceptible to interference from dispersed classes for lower *prop* values. When it comes to producer accuracy, it rose from 57.86% to 98.28% (KNN-5) and from 61.96% to 97.83%.

Crops presented distinguishing values from both classifiers: had a growing pattern in KNN-5 and a decreasing pattern for SVM-OAO. For KNN-5 user accuracy, it grew from 32.57% to 86.25% and for SVM-OAO, its UA varied from 60.34% to 81.77%. Regarding producer accuracy, it varied from 37.25% to 88.03% (KNN-5 PA) and from 78.81% to 88.63% (SVM-OAO PA).

Bare Soil increased all its values. Its user accuracy varied from 52.87% to 89.45% (KNN-5 UA) and from 85.08% to 89.58% (SVM-OAO UA). Its producer accuracy varied from 62.83% to 87.31% (KNN-5) and from 66.52% to 85.69%, presenting similar confusions in this metric.

Water, presented all values converging to 100% elucidating its class heterogeneity.

5.2.4.2.6 Setup 6

This Setup has training samples in $[prop, 100\%]$ and its test samples are in $[prop, 100\%]$. Its resulting charts are in Figure B.23 and Figure B.24. Additionally, an illustration of Setup 6 feature space is presented in Figure A.6. In general, all Water and Forest metric converged to 100% with low range of accuracy for *prop* variations. Another point is that all metrics showed a rising pattern, showing higher accuracy for higher *prop*.

Forest varied from 77.97% to 97.84% (KNN-5 UA) and from 83.48% to 98.68% (SVM-OAO UA). Also, it varied from 74.15% to 98.28% (KNN-5 PA) and from 74.61% to 97.83% (SVM-OAO PA).

Crops, showed a higher variation, presenting variations between 54.28% to 86.25% (KNN-5 UA) and from 48.97% to 83.87% (SVM-OAO UA). Ia also rose from 58.91% to 88.03% (KNN-5 PA) and from 65.68% to 88.63% (SVM-OAO PA)

Regarding Bare Soil, its growth varied from, in user accuracy, from 72.07% to 89.45% (KNN-5 UA) and from 79.04% to 89.58% (SVM-OAO UA). And its producer accuracy varied from from 71.99% to 87.31% (KNN-5 PA) and from 72.68% to 85.69% (SVM-OAO PA).

Water, similarly to the other Setups, presented high accuracy values. Both Forest and Water converged to 100% UA and PA and this did not happen for Crops and Bare Soil. It is possible that this convergence value is related to the fact the segmentation forces classes' pixels values to an average value and that minimised the quality of heterogeneous classes.

5.3 General discussion

The first and perhaps the most important point of this research is the possibility to define the quality of reference data prior to an Remote Sensing image classification. Even though it needs a classified higher-resolution image as auxiliary data, this is a first step towards working on image classifications taking pixel quality into account.

A general comparison between Pixel-based (PIX) and Region-based (REG) baseline classifications lays on the fact that, when using segmentation, there are mode uniform regions, less noisy. This resulted in a lower thematic accuracy. A likely reason for that is on the fact that segmentation forces the region to average pixel values, i.e. mixed within that certain region. This may have affected the image classification analysing purer samples hence not being typical scenarios within regions.

Another point to be addressed in the effect of the baseline classifications on spatial data quality using thematic accuracy and completeness. Both baseline images seemed to define differently the quality of a pixel: PIX seemed to more trustworthy to determine sample quality whereas REG, as part of its characteristics, summarises the quality values.

When comparing thematic accuracy, all Setups for PIX and REG were somehow important to understand to what extent the pixel quality implies the resulting map spatial data quality. A general observation is that using pure test samples ($prop = 100\%$) against mixed training samples implied on higher *kappa* values, indicating that test samples have greater impact on the thematic accuracy assessments. Nonetheless, for thematic accuracy, there was no significant difference between the used classifiers.

Contrarily to thematic accuracy, completeness analysis showed to be dependent on the used classifier, thus indicating that the confusions per class are balanced when thematic accuracy (*kappa* and OA) are used.

Also, even though the four used classes (Forest, Water, Bare Soil and Crops) tend to be spectrally distinct from each other, a minimum sign of correlation between classes seemed to have had drastic effects on user accuracies and producer accuracies, mainly for non-pure pixels ($prop \neq 100\%$). Therefore, analysing class's response to each combination of reference data quality leads us to a better understanding of how it is spectrally divided and how each classifier deals with with.

6 CONCLUSIONS

In a nutshell, we aimed to verify how the selection of reference samples aided by higher resolution imagery could affect the quality assessment of an image classification. The initial step was to build an algorithm to assess the pixel quality of the studied image using statistics of central tendencies. Then, the quality of samples were separated into groups of 5% quality. After that, six selection scenarios (Setups) were tested, combining different variations of training and test quality.

Regarding the algorithm created to define the sample quality, it works successfully and its resulting database is able to be applied for several further analyses for Pattern Recognition applied to Remote Sensing studies. The algorithm is also easy to use and its results are in a GIS familiar format (*shapefile*), hence people non-familiar with programming language can run it.

From the results in the Reference Sample Selection (RSS) approach, the following conclusions are listed:

a) Class representativity is dependent on the sample quality

   From the four used classes, each of them presented different probability distribution function when analysing sample quality, shown by grouped mean, kurtosis and skewness values. Even though the most frequent quality was 100% in all cases, their average value (mass centre) varied and some classes presented broader distribution thus indicating that pure samples may not always describe the studied class.

b) Different baseline classifications represent different perspective for defining the quality of reference data

   Pixel-based baseline classification presented to have their metrics more likely to converge to 100.00 than region-based baseline classification. Therefore, some segmentation characteristics to force region values to their average impacted to determine the quality of reference data, hence pixel-based classification, for this specific scenario, seemed to be reliable.

c) Different classifiers may impact the image classification assessment though not as strongly as the quality of the reference data

   Up to the date of this manuscript, several studies tried to find the best image classifier. However, for this study, difference in the classifier had an effect on the thematic accuracy assessment, but this effect was smaller than the effect of the quality of reference data. This shows that selecting reference samples affects more radically the thematic accuracy studies than the classifier *per se*.

The usage of different classifiers showed to be more relevant for completeness assessment, mainly when there is some heterogeneity in a class, i.e. more dispersed. The classifier ability to differentiate these classes changed drastically for some Setups.

d) The quality of reference samples has an important role in the image classification accuracy assessment

   This fact has been already been proven and it is a usual statement in Remote Sensing but it is confirmed in this manuscript. The purer the pixels, the higher the image classification accuracy.

e) The quality of test samples influence the most on the image classification accuracy assessment

   From all the studied experiments, independent on the situation, we confirmed that this influence is dramatically more significant than the influence of training samples.

To conclude, this study adds a new point of view in Remote Sensing analyses: the importance of the variation of the quality of training and test data as well as the understanding of class representativiness in image classification.

6.1 Recommendations

This manuscript presents an initial study regarding the definition of reference sample quality and analysis of the impact of this quality on the image classification accuracy. As this is an initial study, there are several further questions that raised that may be answered in future research.

The first recommendation is to use more classes to check how Reference Sample Selection responds to more correlated and also heterogeneous classes and verify, or not, the conclusions of this manuscript.

Also, studying the classes' feature spaces, considering class heterogeneity and dispersion in the feature space may aid to further conclusions.

Another point is that more classifiers could be tested to confirm the effect of the classifier on the accuracy assessment.

Regarding combination of Setups, different combinations may be used, using fixed training quality over varied test quality for example so more conclusions over the entire image classification process can be done. There are several Setup combinations and should be testes.

Additionally, image corregistration errors are still common in Remote Sensing and not perfectly corrregistered images may impact the RSS approach. Thus, we recommend to firstly study this error in a controlled environment, testing the registration main parameters separately: shift, rotation, scale and shear.

Finally, this study was completely simulated and theoretical, which implies that this methodology needs to be applied in a real situation, considering different sensors with different spectral responses and a probable corregistration error. Only then, it is possible to determine if this initial research can be applied for the academy.

REFERENCES

ABDOLLAHI, A.; PRADHAN, B. Integrated technique of segmentation and classification methods with connected components analysis for road extraction from orthophoto images. **Expert Systems with Applications**, v. 176, p. 114908–114918, 2021. 10, 18, 19

ALVARES, C. A.; STAPE, J. L.; SENTELHAS, P. C.; GONÇALVES, J. L. D. M.; SPAROVEK, G. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, n. 6, p. 711–728, 2013. 48

ASSIS, L. F. F. G.; FERREIRA, K. R.; VINHAS, L.; MAURANO, L.; ALMEIDA, C.; CARVALHO, A.; RODRIGUES, J.; MACIEL, A.; CAMARGO, C. **TerraBrasilis: a spatial data analytics infrastructure for large-scale thematic mapping**. 2019. Available from: <http://terrabrasilis.dpi.inpe.br/app/dashboard/alerts/legal/amazon/aggregated/#>. 48

BAATZ, M.; HOFFMANN, C.; WILLHAUCK, G. Progressing from object-based to object-oriented image analysis. In: BLASCHKE, T.; LANG, S.; HAY, G. J. (Ed.). **Object-based image analysis: spatial concepts for knowledge-driven remote sensing applications**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. p. 29–42. ISBN 978-3-540-77058-9. Available from: <https://doi.org/10.1007/978-3-540-77058-9_2>. 10

BAATZ, M.; SCHäPE, M. Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. In: STROBL, J.; BLASCHKE, T.; GRIESEBNER, G. (Ed.). **Angewandte geographische Informationsverarbeitung XII**. Berlin, Heidelberg: Wichmann Verlag, 2000. p. 12–23. 10

BELL, P.; BUROVSKI, E.; CAREY, C.; GOMMERS, R.; LARSEN, P. M.; LEE, C.; MA, C.; MCKIBBEN, N.; FORRO, N.; REDDY, T.; WECKESSER, W. **SciPy (1.6.1)**. 2021. Available from: <https://docs.scipy.org/doc/scipy/reference/release.1.6.1.html>. 53

BISHOP, C. M. **Pattern recognition and machine learning (information science and statistics)**. [S.l.]: Springer-Verlag, 2006. 758 p. 29, 30

BLASCHKE, T. Object based image analysis for remote sensing. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 65, n. 1, p. 2–16, 2010. 10

BRANDT, T.; MATHER, P. M. **Classification methods for remotely sensed data**. 2. ed. New York/NY: Taylor & Francis, 2009. 378 p. ISBN 978-1-4200-9072-7. 1, 5, 7, 9, 10, 12, 17, 21, 22, 23, 26, 28, 29

BRAZIL'S NATIONAL INSTITUTE FOR SPACE RESEARCH. **TerraView 5.5.1**. 2020. Available from: <http://www.dpi.inpe.br/terralib5/wiki/doku.php?id=start>. 52, 85

BREIMAN, L. Random forests. **Machine Learning**, v. 45, p. 5–32, 2001. 26, 27

CAIRO, C.; BARBOSA, C.; LOBO, F.; NOVO, E.; CARLOS, F.; MACIEL, D.; JNIOR, R. F.; SILVA, E.; CURTARELLI, V. Hybrid chlorophyll-a algorithm for assessing trophic states of a tropical Brazilian reservoir based on MSI/Sentinel-2 data. **Remote Sensing**, v. 12, n. 1, 2020. ISSN 20724292. 2

CAMPBELL, J. B.; WYNNE, R. H. **Introduction to remote sensing**. 5. ed. [S.l.]: The Guilford Press, 2011. 667 p. ISBN 160918176X. 1

CHANG, N.-B.; BAI, K. **Multisensor data fusion and machine learning for environmental remote sensing**. Boca Raton, Florida. USA.: CRC Press, 2018. 529 p. 21, 22, 23, 24, 26

COCHRAN, W. G. **Sampling techniques**. 3. ed. New York, NY, USA: John Wiley & Sons, 1977. 448 p. 13, 17

CONGALTON, R. G. A review of assessing the accuracy of classifications of remotely sensed data. **Remote Sensing of Environment**, v. 37, p. 35–46, 1991. 35, 37

COUTINHO, A. C.; ALMEIDA, C.; VENTURIERI, A.; ESQUERDO, J. C. D. M.; SILVA, M. **Uso e cobertura da terra nas áreas deflorestadas da Amazônia Legal: TerraClass 2008**. Brasilia, DF, 2013. Available from: <http://ainfo.cnptia.embrapa.br/digital/bitstream/item/87809/1/TerraClass-completo-baixa-pdf.pdf>. 49

ELMES, A.; ALEMOHAMMAD, H.; AVERY, R.; CAYLOR, K.; EASTMAN, J. R.; FISHGOLD, L.; FRIEDL, M. A.; JAIN, M.; KOHLI, D.; BAYAS, J. C. L.; LUNGA, D.; MCCARTY, J. L.; PONTIUS, R. G.; REINMANN, A. B.; ROGAN, J.; SONG, L.; STOYNOVA, H.; YE, S.; YI, Z. F.; ESTES, L. Accounting for training data error in machine learning applied to earth observations. **Remote Sensing**, v. 12, n. 6, p. 1–39, 2020. 1, 2, 9, 12, 13

ESPINDOLA, G. M.; CAMARA, G.; REIS, I. A.; BINS, L. S.; MONTEIRO, A. M. Parameter selection for region-growing image segmentation algorithms using spatial autocorrelation. **International Journal of Remote Sensing**, v. 27, n. 14, p. 3035–3040, 2006. 10

ESRI. **ArcGIS online: shapfiles**. 2020. Available from: <https://doc.arcgis.com/en/arcgis-online/reference/shapefiles.htm>. 57

EUROPEAN SPATIAL AGENCY (ESA). **Copernicus**: [S.l.]: ESA. 2015. 49, 64

______. **ESA Sentinel Online**: MultiSpectral Instrument (MSI) overview. 2015. Available from: <https://earth.esa.int/web/sentinel/technical-guides/sentinel-2-msi/msi-instrument>. Access in: Feb 21 2020. 1, 50, 51

EXÉRCITO BRASILEIRO - DIRETORIA DE SERVIÇO GEOGRÁFICO. **ET-CQDG**: Norma da especificação técnica para controle de qualidade de dados geoespaciais. [S.l.], 2016. 94 p. 3, 27, 34, 35, 36, 37, 65

FOODY, G. M. Status of land cover classification accuracy assessment. **Remote Sensing of Environment**, v. 80, p. 185–201, 2002. ISSN 00344257. 35

FOODY, G. M.; PAL, M.; ROCCHINI, D.; GARZON-LOPEZ, C. X.; BASTIN, L. The sensitivity of mapping methods to reference data quality: training supervised image classifications with imperfect reference data. **ISPRS International Journal of Geo-Information**, v. 5, n. 11, 2016. ISSN 22209964. 2

GILLIES, S. et al. **Rasterio: geospatial raster I/O for Python programmers**. 2013–. Available from: <https://github.com/mapbox/rasterio>. 53

GUPTILL, S.; MORRISON, J. **Elements of spatial data quality**. [S.l.]: International Cartographic Association Commission on Spatial Data Quality, 1995. 33

HADDAD, K.; RAHMAN, A.; A ZAMAN M., S. S. Applicability of monte carlo cross validation technique for model development and validation using generalised least squares regression. **Journal of Hydrology**, v. 482, p. 119–128, 2013. 32

HARALICK, R.; SHANMUGAM, K.; DINSTEIN, I. Textural Features for Image Classification. **Syst. Man Cybern. IEEE Trans.**, v. 6, p. 610–621, 1973. 11

HARRIS, C. R.; MILLMAN, K. J.; WALT, S. J. van der; GOMMERS, R.; VIRTANEN, P.; COURNAPEAU, D.; WIESER, E.; TAYLOR, J.; BERG, S.; SMITH, N. J.; KERN, R.; PICUS, M.; HOYER, S.; KERKWIJK, M. H. van; BRETT, M.; HALDANE, A.; R'ıO, J. F. del; WIEBE, M.; PETERSON, P.; G'eRARD-MARCHANT, P.; SHEPPARD, K.; REDDY, T.; WECKESSER, W.; ABBASI, H.; GOHLKE, C.; OLIPHANT, T. E. Array programming with NumPy. **Nature**, v. 585, n. 7825, p. 357–362, sep. 2020. 53, 56

HASTIE, T.; TIBDHIRANI, R.; FRIEDMAN, J. **The elements of statistical learning: data mining, inference, and prediction**. 2. ed. [S.l.]: Springer, 2009. 2, 26, 27, 29, 30, 32

HOSSAIN, M. D.; CHEN, D. Segmentation for Object-Based Image Analysis (OBIA): a review of algorithms and challenges from remote sensing perspective. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 150, n. November 2018, p. 115–134, 2019. 10

HUDSON, R. D.; HUDSON, J. W. The military applications of remote sensing by infrared. **Proceedings of the IEEE**, n. 1, p. 104–128. ISSN 15582256. 1

HUNTER, J. D. Matplotlib: a 2d graphics environment. **Computing in Science & Engineering**, v. 9, n. 3, p. 90–95, 2007. 53

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 19157:2013**: Geographic information - data quality. [S.l.], 2013. 146 p. Available from: <https://www.iso.org/obp/ui/#iso:std:iso:19157:ed-1:v1:en>. 27, 33, 34, 35, 36, 37

INTITUTO BRASILEIRO DE GEOGRAFIA E ESTATíSTICA. **Malha municipal**. Rio de Janeiro: IBGE, 2020. 49

JAMES, G.; WITTEN, D.; HASTIE, T.; TIBSHIRANI, R. **An Introduction to statistical learning**. 2. ed. [S.l.]: Springer Texts in Statistics, 2021. 612 p. 29, 30, 31

JIN, H.; STEHMAN, S. V.; MOUNTRAKIS, G. Assessing the impact of training sample selection on accuracy of an urban classification: a case study in denver, colorado. **International Journal of Remote Sensing**, v. 35, n. 6, p. 2067–2081, 2014. ISSN 13665901. 2

JORDAHL, K. **GeoPandas: Python tools for geographic data**. **Available from:** https://github.com/geopandas/geopandas, 2014. 53

KÖRTHING, T. S. **GeoDMA: a toolbox for integrating data mining with object-based and multi-temporal analysis of satellite**. Thesis (PhD in Remote Sensing) — Brazil National Institute for Space Research, São José dos Campos, Sao Paulo, Brazil, 2012. Available from: <http://mtc-m16d.sid.inpe.br/rep/8JMKD3MGP7W/3CCH86S?ibiurl.backgroundlanguage=en>. 10, 25

KÖRTHING, T. S.; FONSECA, L. M. G.; CÂMARA, G. GeoDMA - geographic data mining analyst. **Computers & Geosciences**, v. 57, 2013. Available from: <http://www.dpi.inpe.br/geodma>. 25, 26

KUHN, M.; JOHNSON, K. **Applied predictive modeling with applications in R**. [s.n.], 2013. 615 p. Available from: <http://appliedpredictivemodeling.com/s/Applied_Predictive_Modeling_in_R.pdf>. 25, 26, 32

LARY, D. J.; ZEWDIE, G. K.; LIU, X.; WU, D.; LEVETIN, E.; ALLEE, R. J.; MALAKAR, N.; WALKER, A.; MUSSA, H.; MANNINO, A.; AURIN, D. Machine learning applications for Earth observation. In: MATHIEU P. P.; AUBRECHT, C. E. (Ed.). **Earth observation, open science and innovation**. Cham, Switzerland: Springer International Publishing, 2018. p. 165–218. ISBN 9783319656335. 1

LAUNIUS, R.; LOGSDON, J.; SMITH, R. W. **Reconsidering Sputnik**: forty years since the soviet satellite. [S.l.]: Routledge, 2002. 1

LYONS, M. B.; KEITH, D. A.; PHINN, S. R.; MASON, T. J.; ELITH, J. A comparison of resampling methods for remote sensing classification and accuracy assessment. **Remote Sensing of Environment**, v. 208, p. 145–153, 2018. 29, 32

MACIEL, D.; NOVO, E.; CARVALHO, L. S. de; BARBOSA, C.; FLORES JÚNIOR, R.; LOBO, F. d. L. Retrieving total and inorganic suspended sediments in Amazon floodplain lakes: a multisensor approach. **Remote Sensing**, v. 11, n. 15, 2019. ISSN 20724292. 2

MELLOR, A.; BOUKIR, S.; HAYWOOD, A.; JONES, S. Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 105, p. 155–168, 2015. ISSN 09242716. Available from:
<http://dx.doi.org/10.1016/j.isprsjprs.2015.03.014>. 2

NATIONAL INSTITUTE FOR SPACE RESEARCH (INPE). **Programa de monitoramento da Amazônia e demais biomas**: Dematamento - amazônia legal. São José dos Campos, SP: INPE, 2019. Available from: <http://terrabrasilis.dpi.inpe.br/downloads/>. Access in: 13 Jan 2022. 48

NOVO, E. M. L. d. M. **Sensoriamento remoto: princípios e aplicações**. 4. ed. São Paulo, SP, Brazil: Blucher, 2010. 5, 10

OLOFSSON, P.; FOODY, G. M.; HEROLD, M.; STEHMAN, S. V.; WOODCOCK, C. E.; WULDER, M. A. Good practices for estimating area and assessing accuracy of land change. **Remote Sensing of Environment**, Elsevier Inc., v. 148, p. 42–57, 2014. 12, 13, 14, 15, 16, 17

PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, E. Scikit-learn: machine learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011. 53, 56

PICARD, R. R.; COOK, R. D. Cross-validation of regression models. **Journal of the American Statistical Association**, v. 79, n. 387, p. 575–583, 1984. 32

PLOTY TECHNOLOGIES INC. **Collaborative data science**. Montreal, QC, Canada: Plotly Technologies, 2015. Available from: <https://plot.ly>. 53

QGIS DEVELOPMENT TEAM. **QGIS Geographic Information System**. [S.l.], 2009. Available from: <http://qgis.osgeo.org>. 52

REIS, M. S.; DUTRA, L. V.; ESCADA, M. I. S.; SANT'ANNA, S. J. S. Avoiding invalid transitions in land cover trajectory classification with a compound maximum a posteriori approach. **IEEE Access**, v. 8, p. 98787–98799, 2020. 49, 50

REIS, M. S.; ESCADA, M. I. S.; DUTRA, L. V.; SANT'ANNA, S. J.; VOGT, N. D. Towards a reproducible LULC hierarchical class legend for use in the Southwest of Pará State, Brazil: a comparison with remote sensing data-driven hierarchies. **Land**, v. 7, n. 2, 2018. 49, 50

REIS, M. S.; ESCADA, M. I. S.; SANT'ANNA, S. J. S.; DUTRA, L. V. Land use and land cover trajectory classification and analysis methods in the Amazon: Implications for forest regeneration studies. **Revista Brasileira de Cartografia**, v. 72, 2020. 49

REIS, S. M. **Detecção de mudanças de uso e cobertura da terra utilizando dados óticos e de micro-ondas em uma região da Amazônia Brasileira**. 290 p. Dissertation (Masters in Remote Sensing) — National Institute for Space Research (INPE), São José dos Campos (SP), 2014. Available from: <http://urlib.net/rep/8JMKD3MGP6W34M/3PSMBRH?ibiurl.backgroundlanguage=pt-BR>. 49

RICHARDS, J. A.; XIUPING, J. **Remote sensing digital image analysis: an introduction**. 4. ed. Berlin, Germany: Springer, 2006. 1, 2, 5, 6, 7, 9, 10, 18, 20, 21, 22, 25, 28

RITTER, N.; RITH, M. **GeoTIFF format specification**. 2000. Available from: <http://geotiff.maptools.org/spec/geotiffhome.html>. 57

ROSSUM, G. V.; DRAKE, F. L. **Python 3 reference manual**. Scotts Valley, CA: CreateSpace, 2009. ISBN 1441412697. 52, 56

SANTOS, A. d. P. dos. **Cartographic quality control: methodologies for evaluation of positional accuracy in spatial data**. 188 p. Thesis (PhD in Civil Engineering — Universidade Federal de Vicosa (UFV), Viçosa (MG), 2015. 33

SHANMUGAM, R.; CHATTAMVELLI, R. In: **Statistics for scientists and engineers**. [S.l.: s.n.]. 42

SHAO, J. Linear model selection by cross-validation. **Journal of the American Statistical Association**, v. 88, n. 422, p. 486–494, 1993. 32

SHIMABUKURO, Y. E.; PONZONI, F. J. **Spectral mixture for remote sensing: linear model and applicaitons**. Cham, Switzerland: Springer International Publishing, 2019. 80 p. ISBN 9783030020163. Available from: <http://dx.doi.org/10.1007/978-3-030-02017-0>. Access in: 14 Sept. 2021. 2

SOARES, M. D.; DUTRA, L. V.; COSTA, G. A. O. P. da; FEITOSA, R. Q.; NEGRI, R. G.; DIAZ, P. M. A meta-methodology for improving land cover and land use classification with SAR imagery. **Remote Sensing**, v. 12, n. 6, p. 1–18, 2020. 49

SOTHE, C. **Mapping successional forest stages and tree species in subtropical areas integrating uav-based photogrammetric point cloud and hyperspectral data : comparison of machine and deep learning**

algorithms. Thesis (PhD in Remote Sensing) — National Institute for Space Research (INPE), São José dos Campos, São Paulo, Brazil, 2019. 27

STEHMAN, S. V. Basic probability sampling designs for thematic map accuracy assessment. **International Journal of Remote Sensing**, v. 20, n. 12, p. 2423–2441, 1999. 2, 12, 13, 14, 15, 16

______. Sampling designs for accuracy assessment of land cover. **International Journal of Remote Sensing**, v. 30, n. 20, p. 5243–5272, 2009. 2, 9, 12, 13, 14, 15, 16, 17

STEHMAN, S. V.; CZAPLEWSKI, R. L. Design and analysis for thematic map accuracy assessment: fundamental principles. **Remote Sensing of Environment**, v. 64, n. 3, p. 331–344, 1998. 2

STEHMAN, S. V.; FOODY, G. M. Key issues in rigorous accuracy assessment of land cover products. **Remote Sensing of Environment**, v. 231, p. 111199, 2019. 2, 12, 13, 15

THE PANDAS DEVELOPMENT TEAM. **pandas-dev/pandas: Pandas**. Zenodo, Feb 2020. Available from: <https://doi.org/10.5281/zenodo.3509134>. 53

THEODORIDIS, S.; KOUTROUMBAS, K. **Pattern recognition**. 4. ed. [S.l.]: Academic Press, 2009. 2, 5, 6, 7, 18, 21, 23, 24, 25, 26, 29, 30, 32

Trimble Germany GmbH. **Trimble Documentation eCognition Developer 9.0.1 Reference Book**. [S.l.], 2014. 10, 11, 52, 85

VAN OORT, P. **Spatial data quality: from description to application**. 2006. 125 p. Thesis (PhD in Production Ecology and Resource Conservation) — Wageningen Universiteit, Wageninger, Netherlands, 2006. Available from: <https://library.wur.nl/WebQuery/wurpubs/346112>. Access in: 28 Nov 2021. 37

VRIGAZOVA, B. The proportion for splitting data into training and test set for the bootstrap in classification problems. **Business Systems Research**, v. 12, n. 1, p. 228–242, 2021. 31

WARNER, T. A.; NELLIS, M. D.; FOODY, G. M. **The SAGE handbook of remote sensing**. [S.l.]: SAGE Publications, 2009. 12

GLOSSARY

Accuracy assessment analysis of a thematic map, combining two elements of spatial data quality: Thematic Accuracy and Completeness 83, *see* spatial data quality

Baseline image classified higher-resolution image used as basis for Reference Sample Selection approach 39, 110

Class representativiness how a certain class is distributed considering its average quality and standard deviation. 58, 67, 68, 76, 92

Classification model models used draw conclusions from the input values given from the training samples. It is supposed to predict the class category/label for new data. 5, 20

Completeness part of the analysis of a thematic map. It is an element of Spatial data quality. It is divided into Commission and Omission errors (%), or Producer and User Accuracy (%) 65, 81, 82, 84, 89, 105, 106, 110, *see* spatial data quality

Feature space in Remote Sensing, it is a n-dimensional space according to the number of studied bands (features) where each pixel is a vector of greyscale values 5–7, 18–24, 74–76, 81–84, 105–109

Producer accuracy PA, part of the analysis of a thematic map. It is a metric of Completeness, it is related to Omission Error (OE), defined by $PA = 1 - OE$. It is related to the absence of data in the studied class, resulting in incompleteness. 37, 81, 83, 85, 105–110

Reference data qata used as reference for a certain process, which can be the classification model or accuracy assessment in this study. Also called as reference sample, labelled samples or even labelled data. 1, 7, 9, 12, 13, 28, 29, 35, 39–44, 47, 56, 57, 60, 63, 67, 73–76, 81, 85, 97, 98, 105, 110–112

Spatial data quality a degree of data excellency to satisfy a given objective 33, 84, 85, 110, 123, 124

Test sample data used for assessing thematic accuracy of the classification result. Also called as classification validation reference. xix, 2, 3, 7, 28, 42–44, 56, 60, 73, 74, 81–86, 97–100, 105–109

Thematic accuracy part of the analysis of a thematic map. It is an element of Spatial data quality. It is represented by overall accuracy (%) or *kappa* index 81, 83, 86, 87, 97–100, 105, 109, 110, *see* spatial data quality

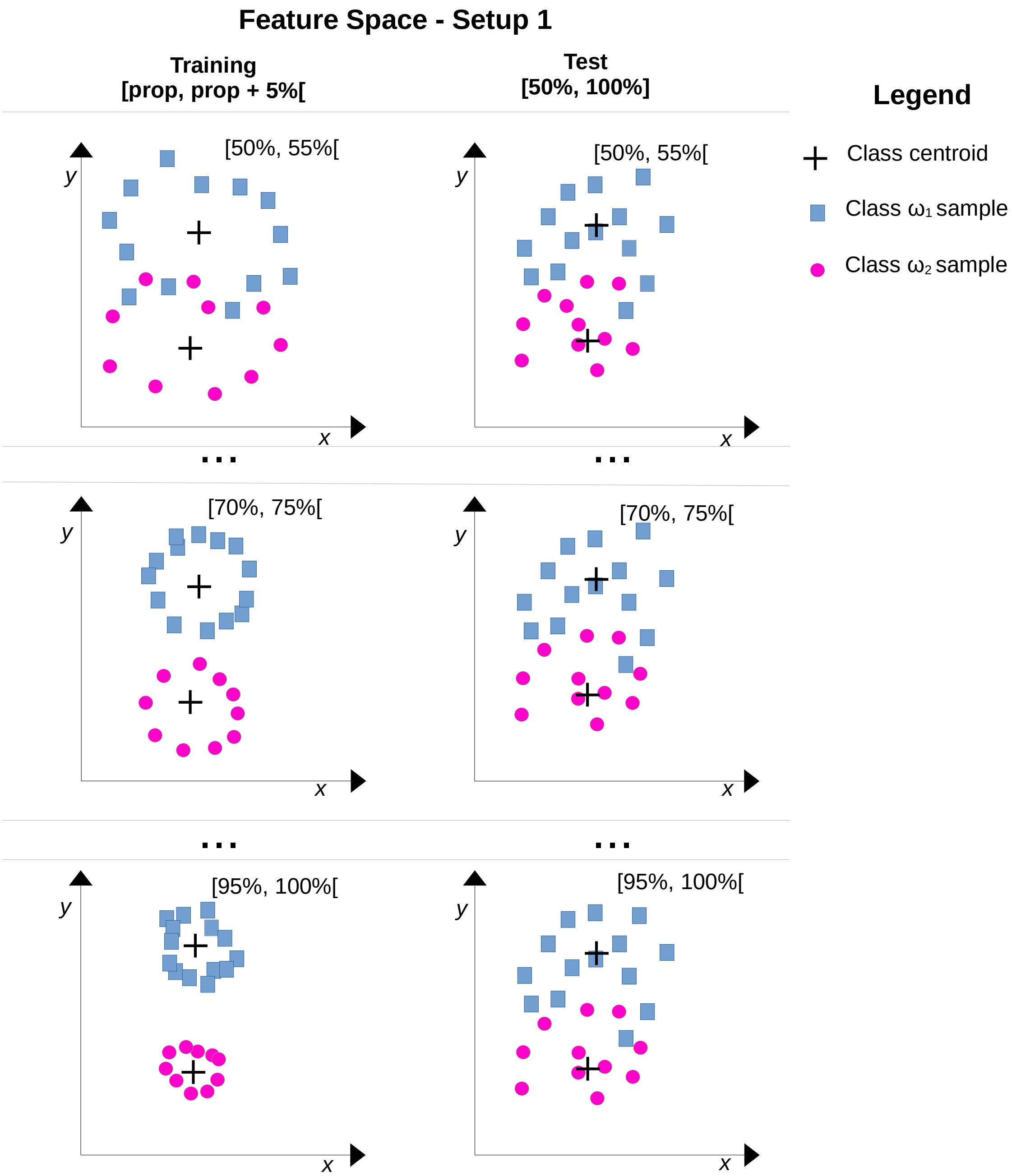
Training sample set of pixels or segmented region used define the parameters for the classification model. Also called as classification training reference. xix, 13, 19, 20, 30, 44, 56, 59, 73–75, 82–86, 98–100, 105–109

User accuracy UA, part of the analysis of a thematic map. It is a metric for Completeness, it is related to Commission Error (CE), defined by $UA = 1 - CE$. It is related to the excess of data in the studied class, resulting in overcompleteness. 37, 65, 81, 83, 84, 89, 105–110

APPENDIX A - SIMULATED FEATURE SPACE PER SETUP

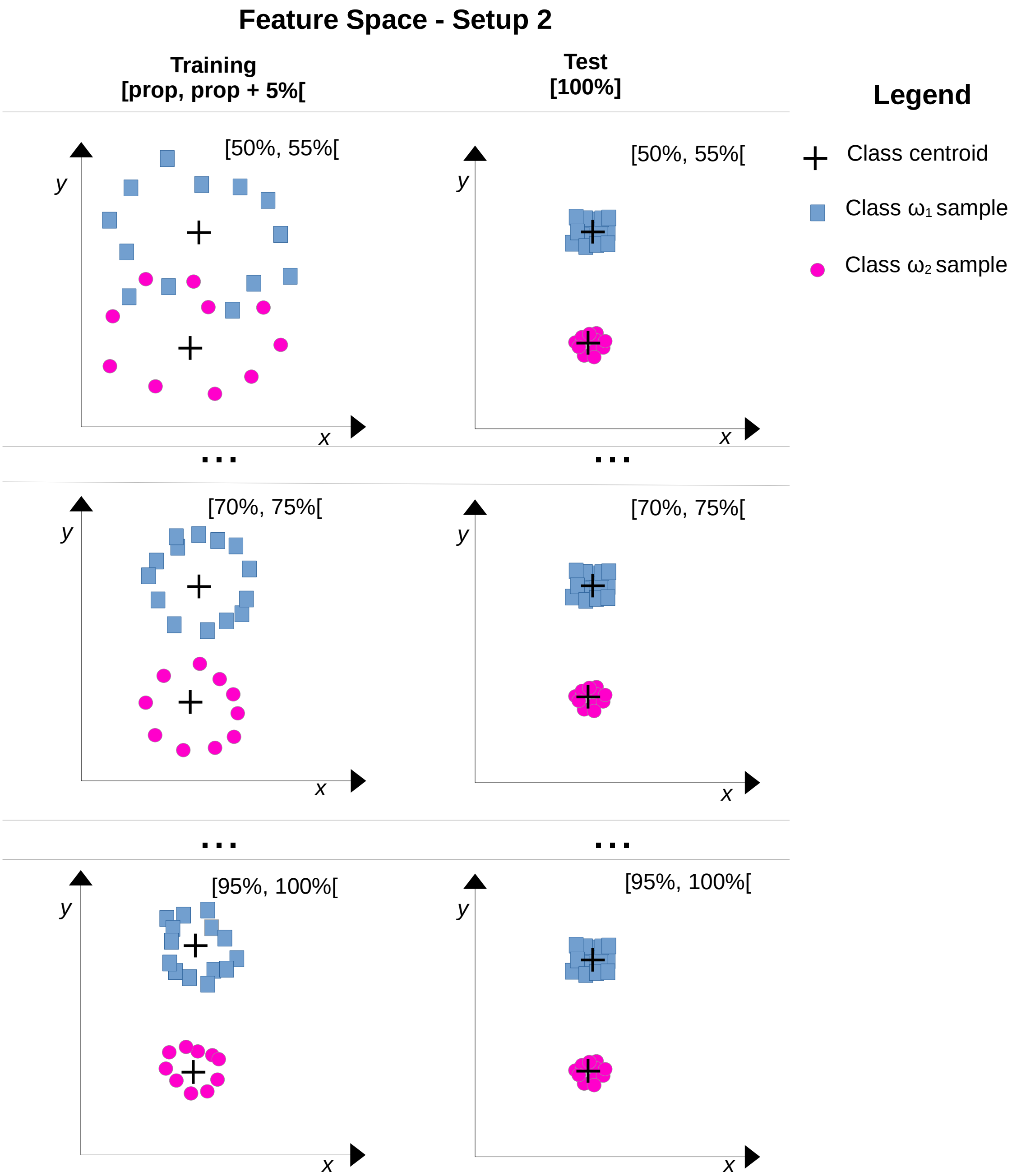
This appendix presented the simulated features spaces of all six Setups to complement the understanding in Chapter 3.

Figure A.1 - Simulated Feature Space for Setup 1.



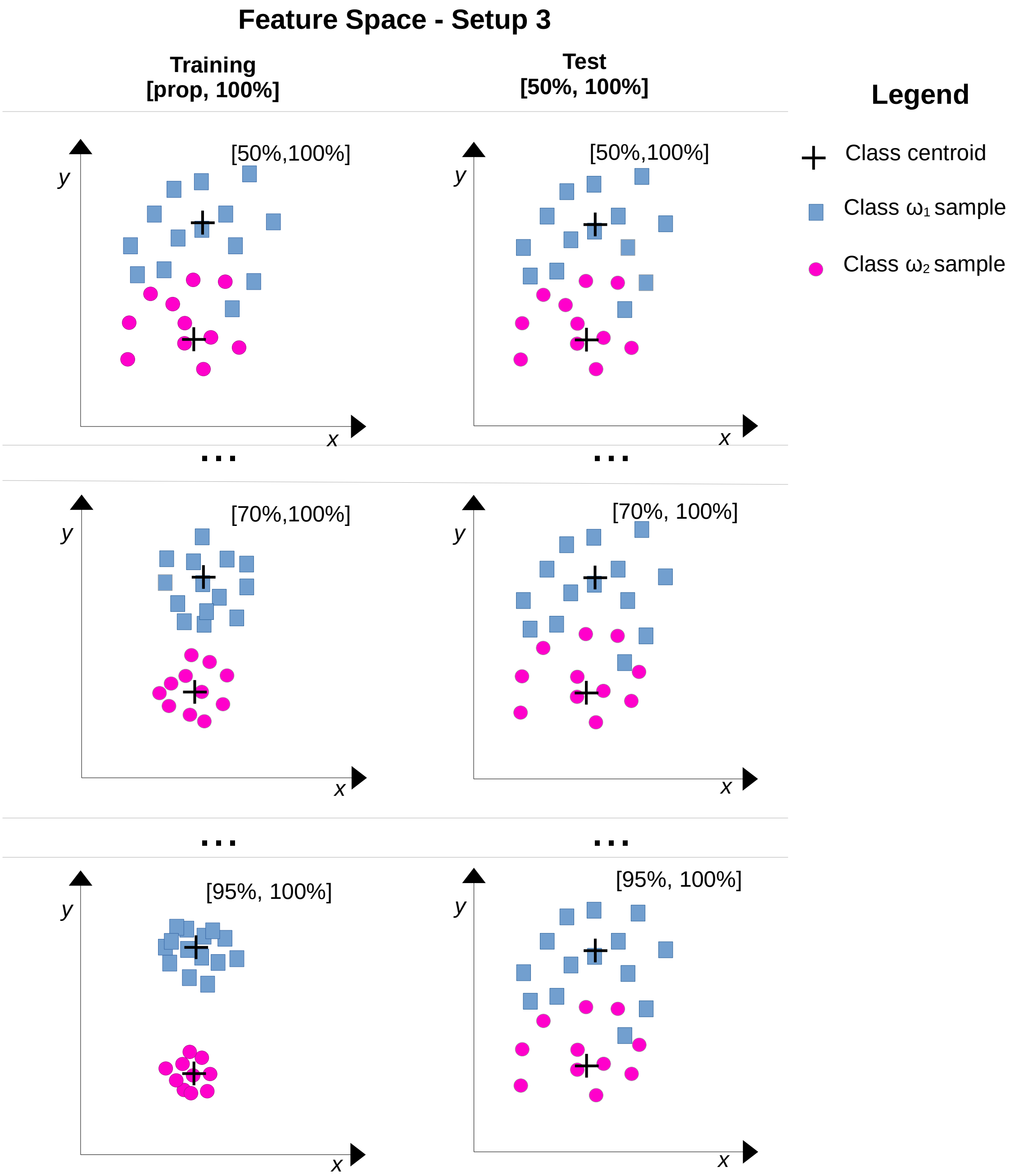
SOURCE: Author.

Figure A.2 - Simulated Feature Space for Setup 2.



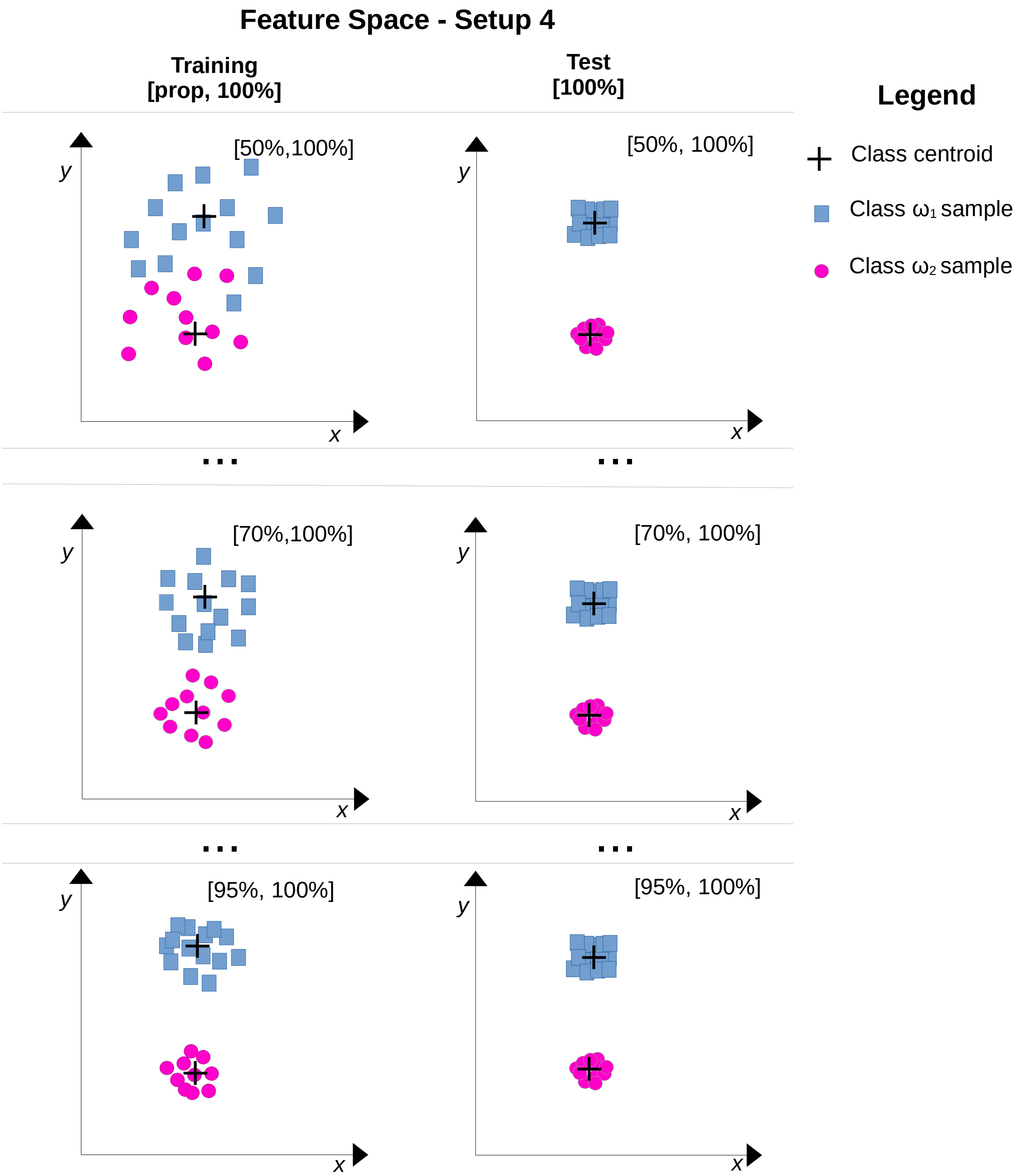
SOURCE: Author.

Figure A.3 - Simulated Feature Space for Setup 3.



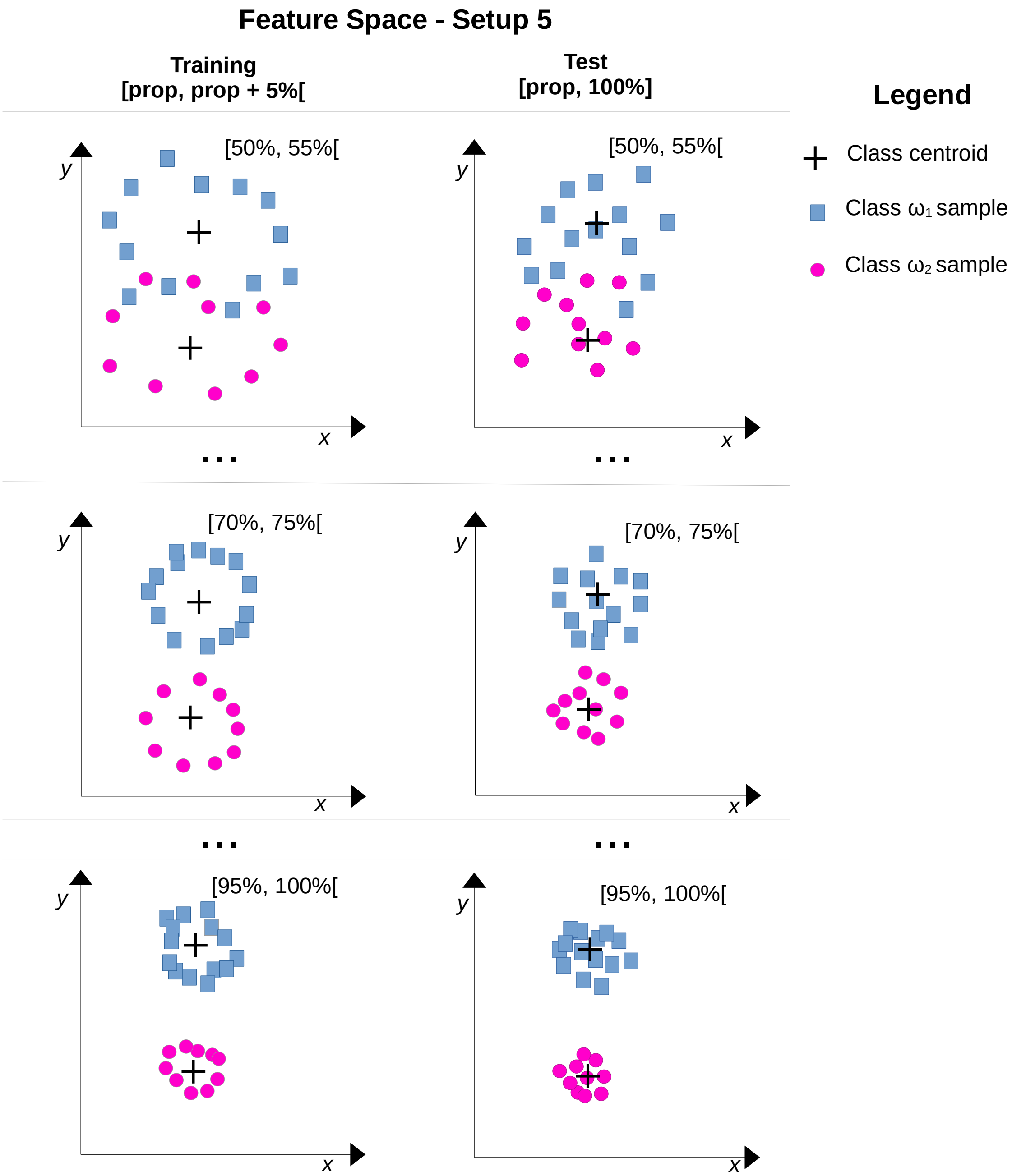
SOURCE: Author.

Figure A.4 - Simulated Feature Space for Setup 4.



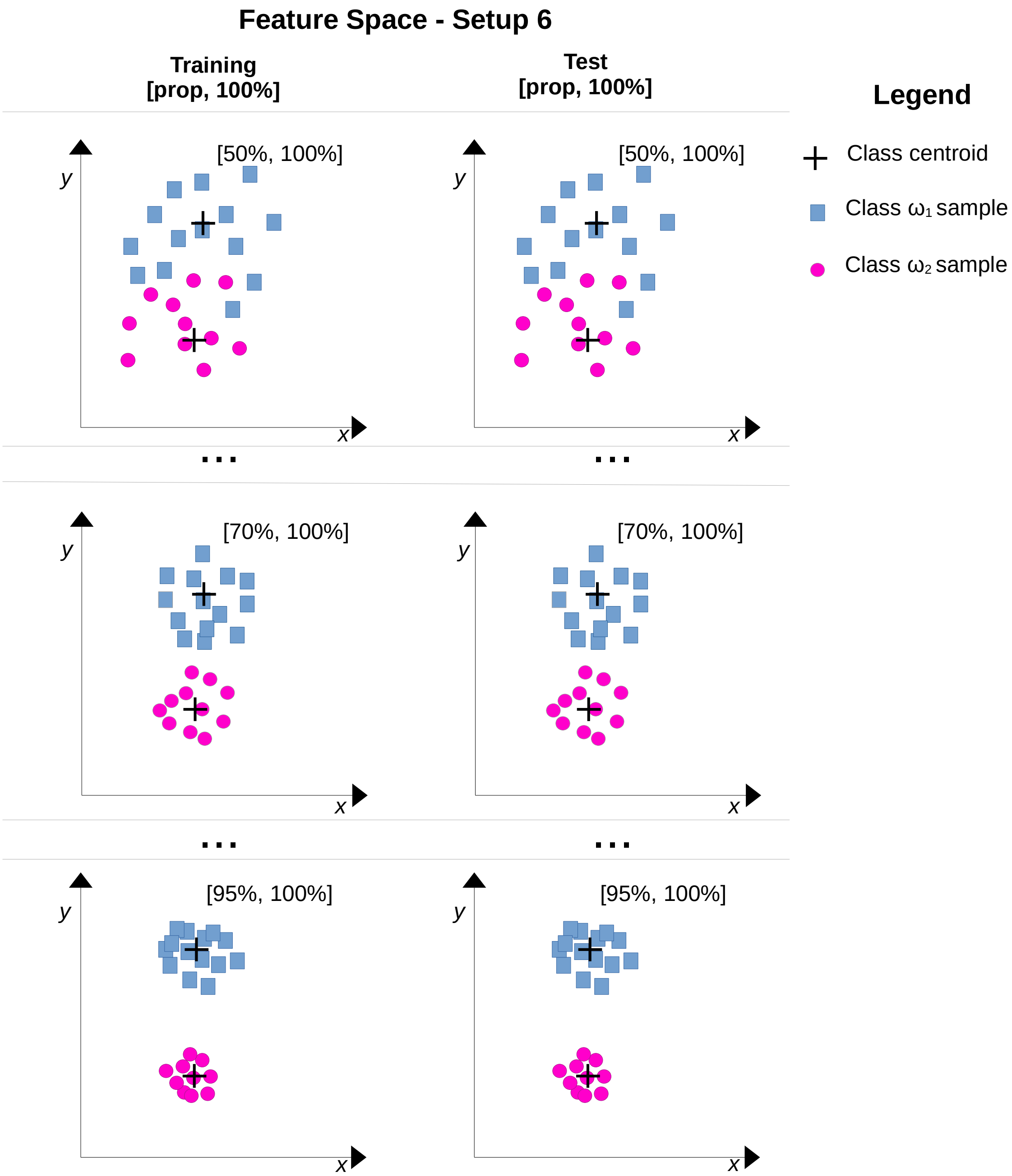
SOURCE: Author.

Figure A.5 - Simulated Feature Space for Setup 5.



SOURCE: Author.

Figure A.6 - Simulated Feature Space for Setup 6.



SOURCE: Author.

APPENDIX B - THEMATIC ACCURACY AND COMPLETENESS FROM REFERENCE SAMPLE SELECTION METHODOLOGY

This appendix presents the charts regarding the elements of Thematic Accuracy and Completeness from Spatial Data Quality. Each Figure presents the Thematic Accuracy represented by the *kappa* index and Completeness represented by User Accuracy (UA) and Producer Accuracy (PA).

B.1 RSS part II - classification results for Pixel-based baseline classification

In this part, the results for data with the Pixel-based Baseline classification spatial data quality are presented. Firstly the Graphic results are presented (Appendix B.1.1) leading to a visual interpretation of the data. Afterwards the tables with exact data are presented (Appendix B.1.2).

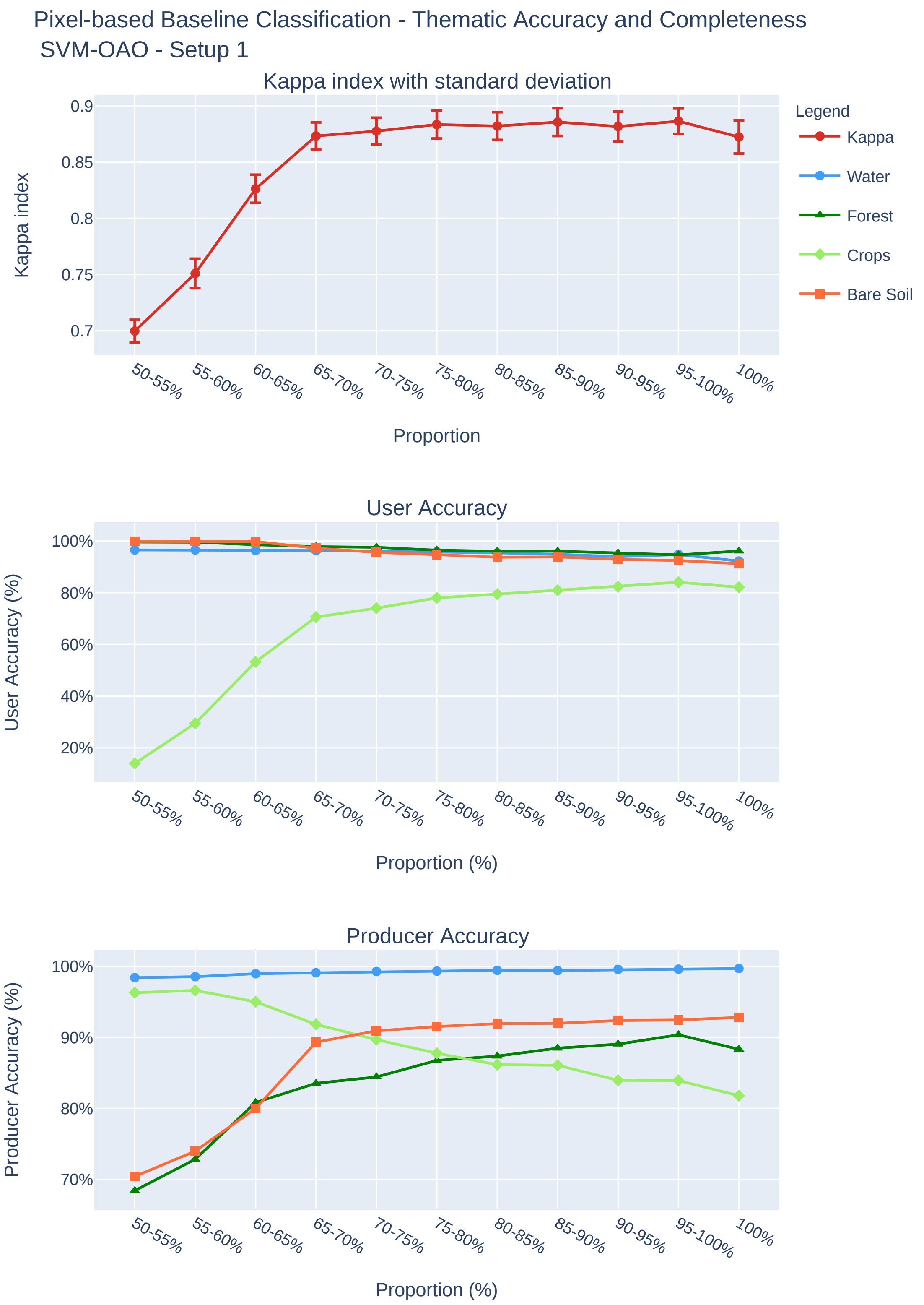
B.1.1 Graphic results

The graphic results present graphs with the mean *kappa* index and the error bars representing their respective standard deviation (std) in the upper area. The middle area presents the mean User Accuracy (UA) whilst the lower area presents the Producer Accuracy (PA).

We note that the vertical scale of all graphs may change according to minimum and maximum value of the presented information.

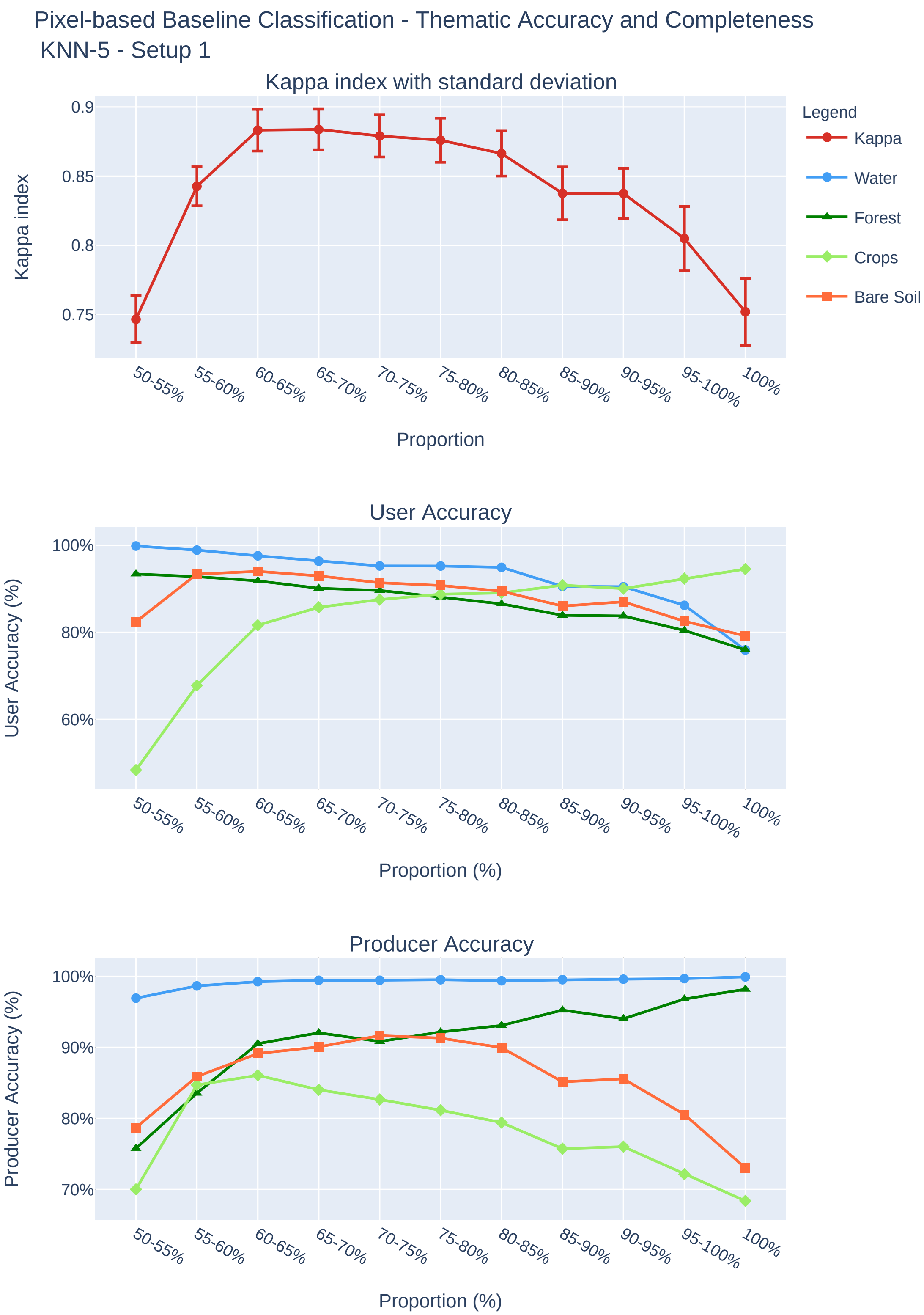
All figures in this section refer to Section 5.1.4, in page 73.

Figure B.1 - Pixel-based Baseline Classification - Accuracy Assessment for SVM-OAO for Setup 1.



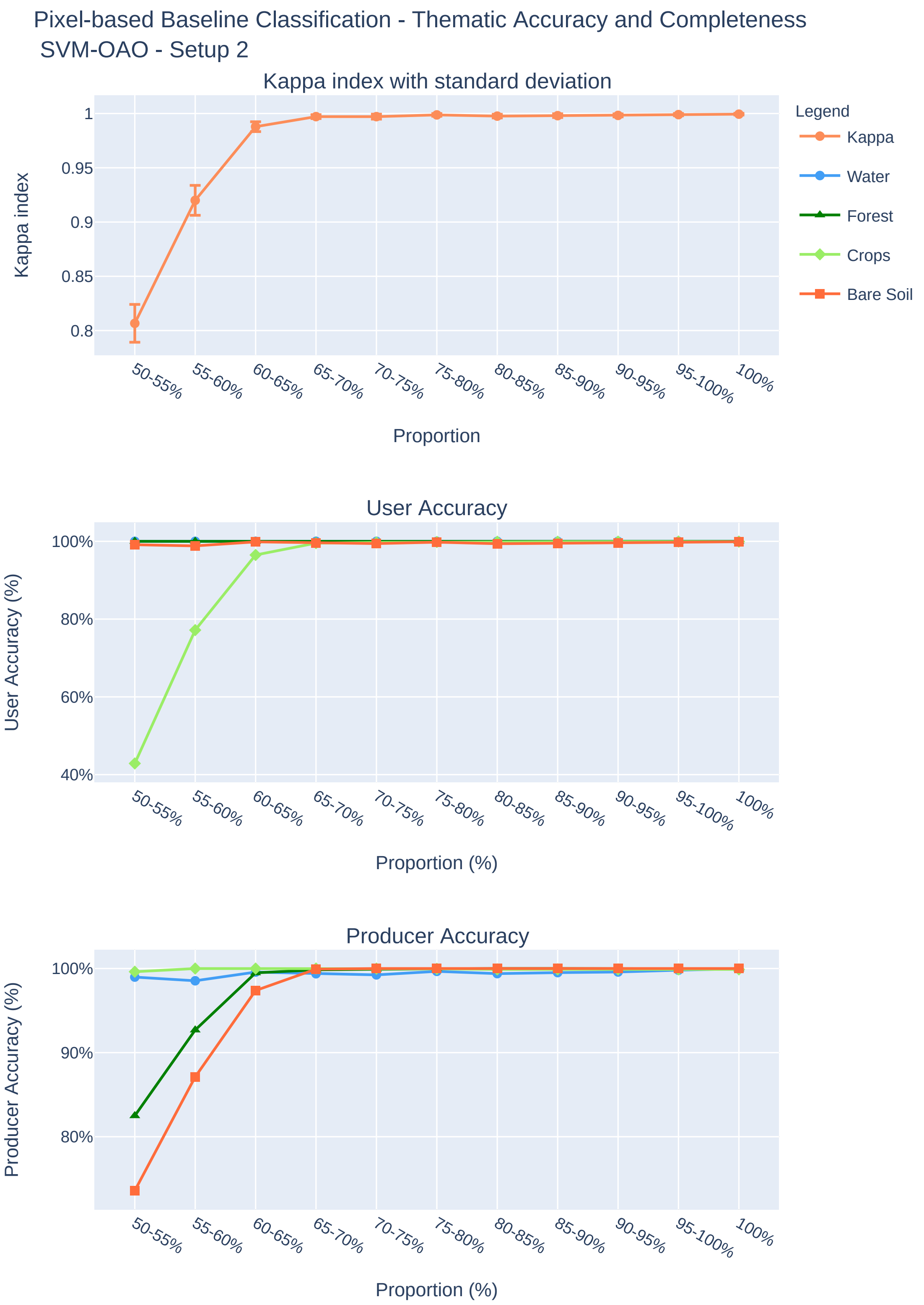
SOURCE: Author.

Figure B.2 - Pixel-based Baseline Classification - Accuracy Assessment for KNN-5 for Setup 1.



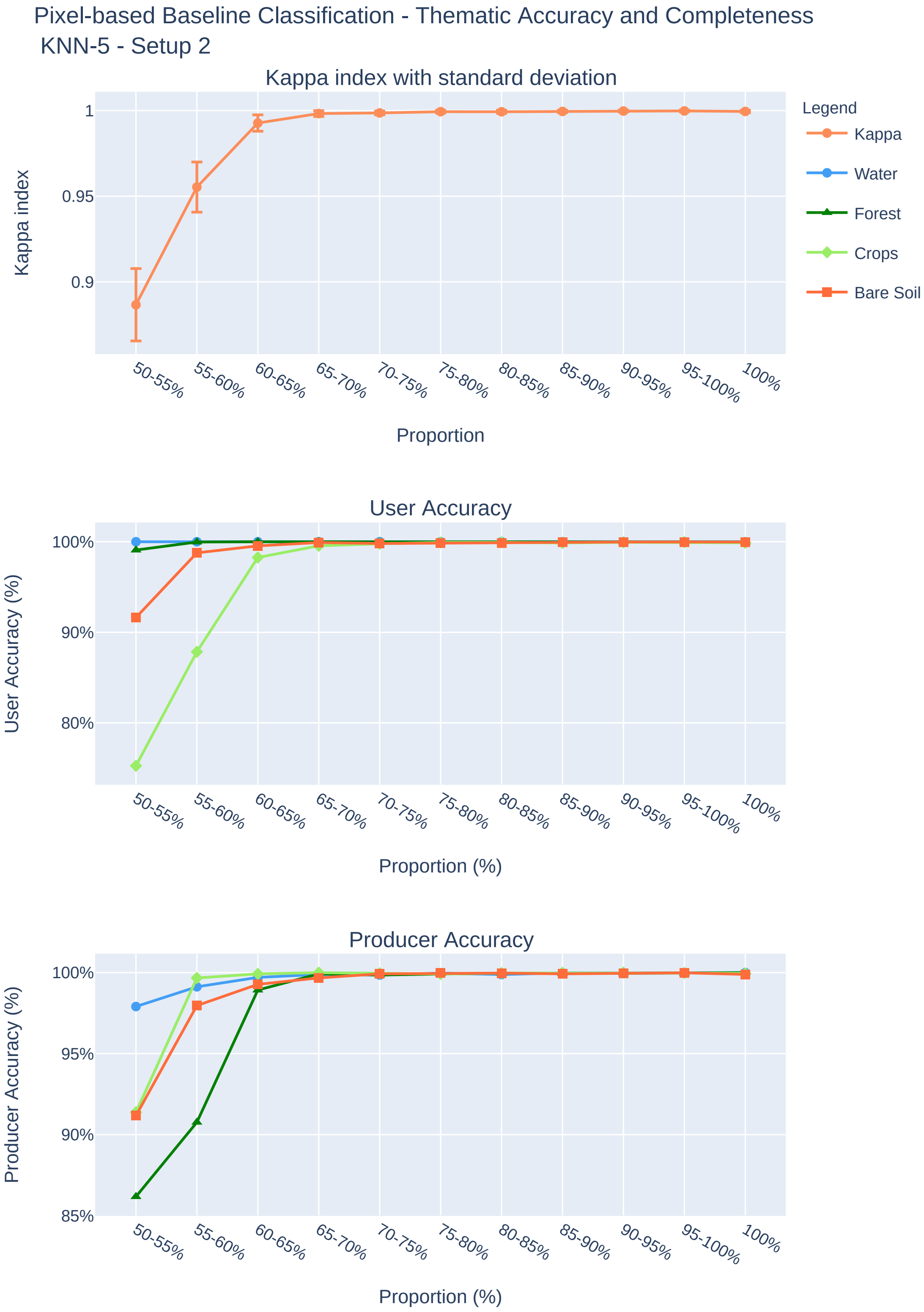
SOURCE: Author.

Figure B.3 - Pixel-based Baseline Classification - Accuracy Assessment for SVM-OAO for Setup 2.



SOURCE: Author.

Figure B.4 - Pixel-based Baseline Classification - Accuracy Assessment for KNN-5 for Setup 2.

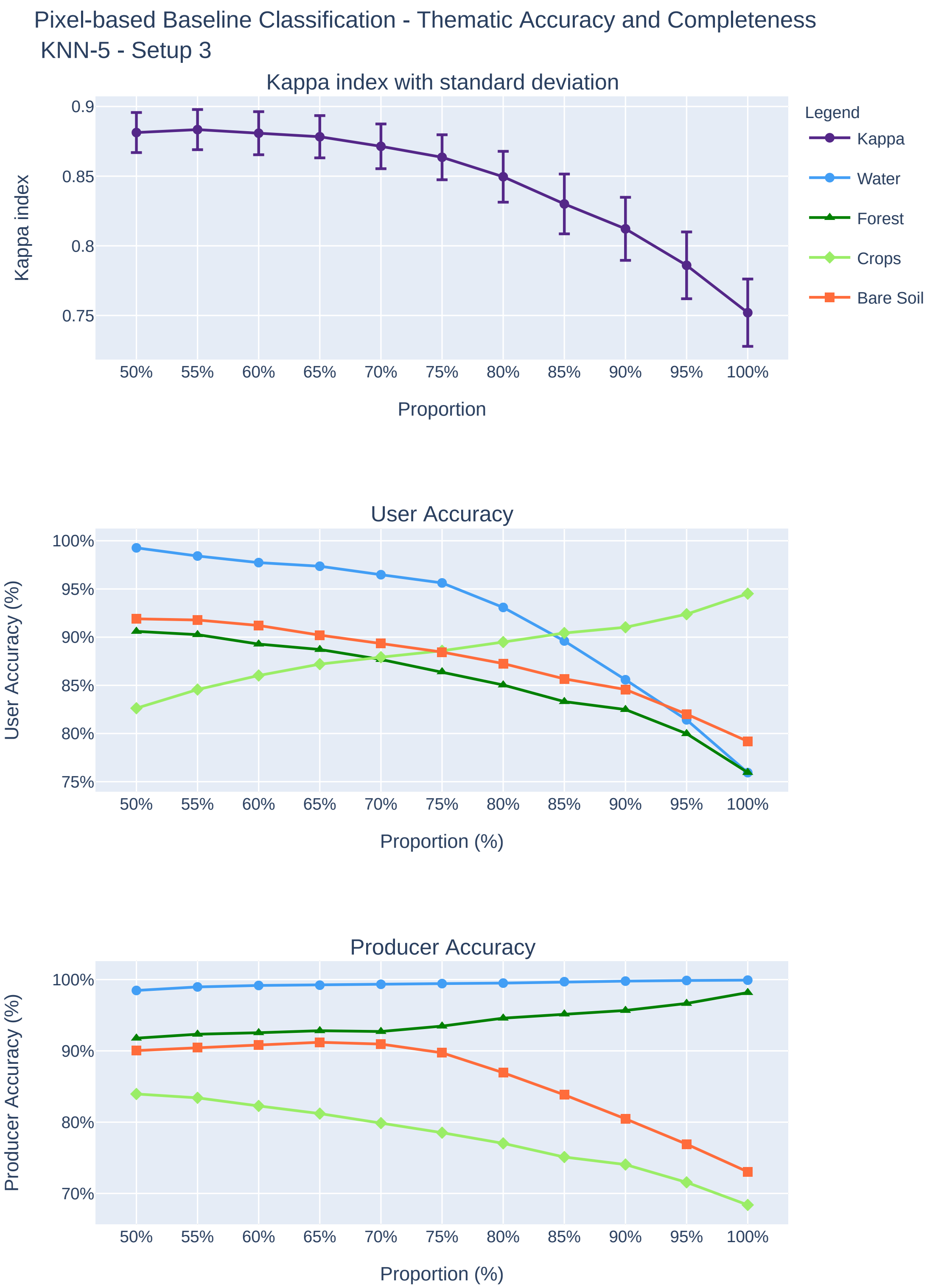


SOURCE: Author.

Figure B.5 - Pixel-based Baseline Classification - Accuracy Assessment for SVM-OAO for Setup 3.

SOURCE: Author.

Figure B.6 - Pixel-based Baseline Classification - Accuracy Assessment for KNN-5 for Setup 3.



SOURCE: Author.

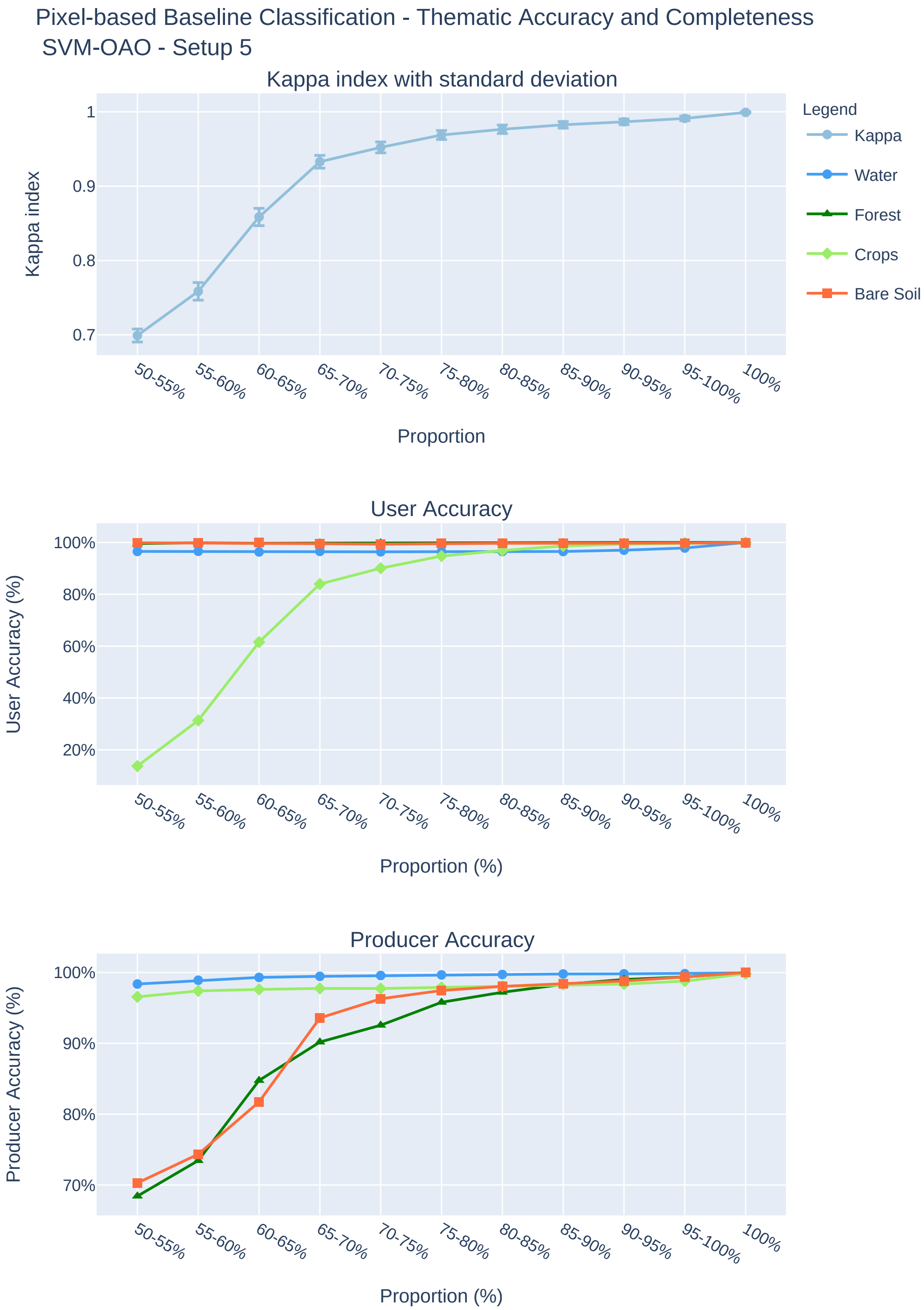
Figure B.7 - Pixel-based Baseline Classification - Accuracy Assessment for SVM-OAO for Setup 4.

SOURCE: Author.

Figure B.8 - Pixel-based Baseline Classification - Accuracy Assessment for KNN-5 for Setup 4.

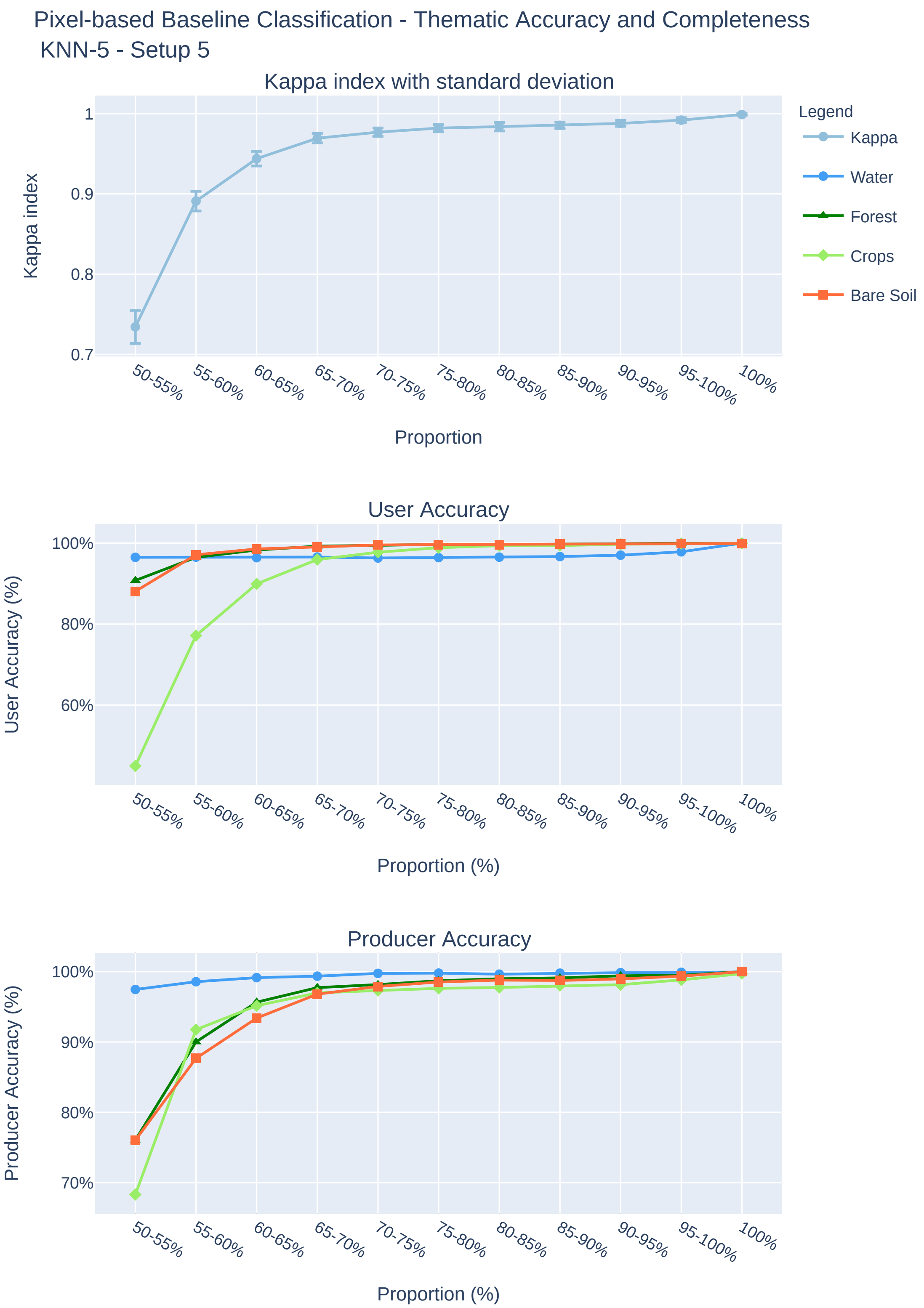
SOURCE: Author.

Figure B.9 - Pixel-based Baseline Classification - Accuracy Assessment for SVM-OAO for Setup 5.



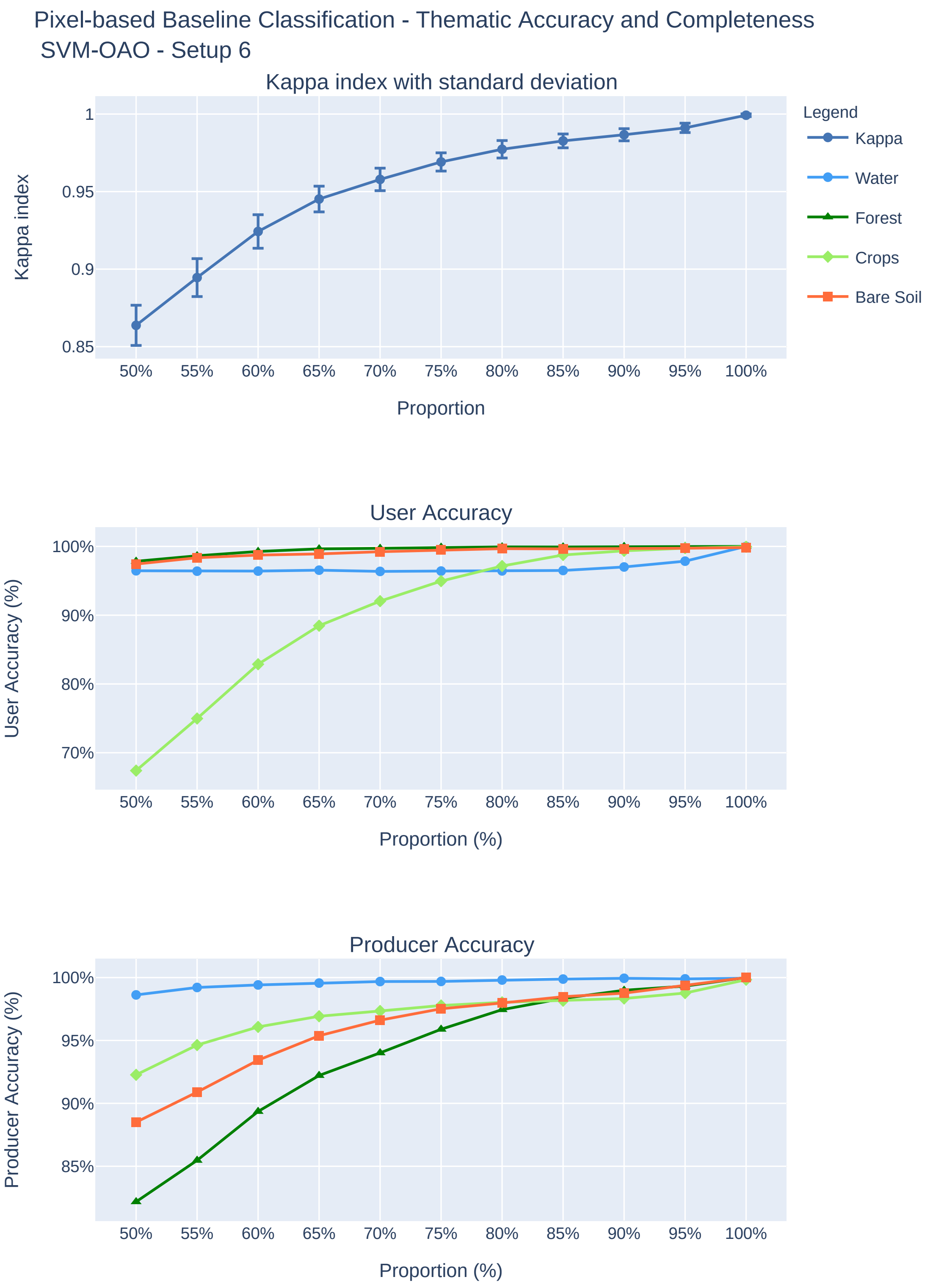
SOURCE: Author.

Figure B.10 - Pixel-based Baseline Classification - Accuracy Assessment for KNN-5 for Setup 5.



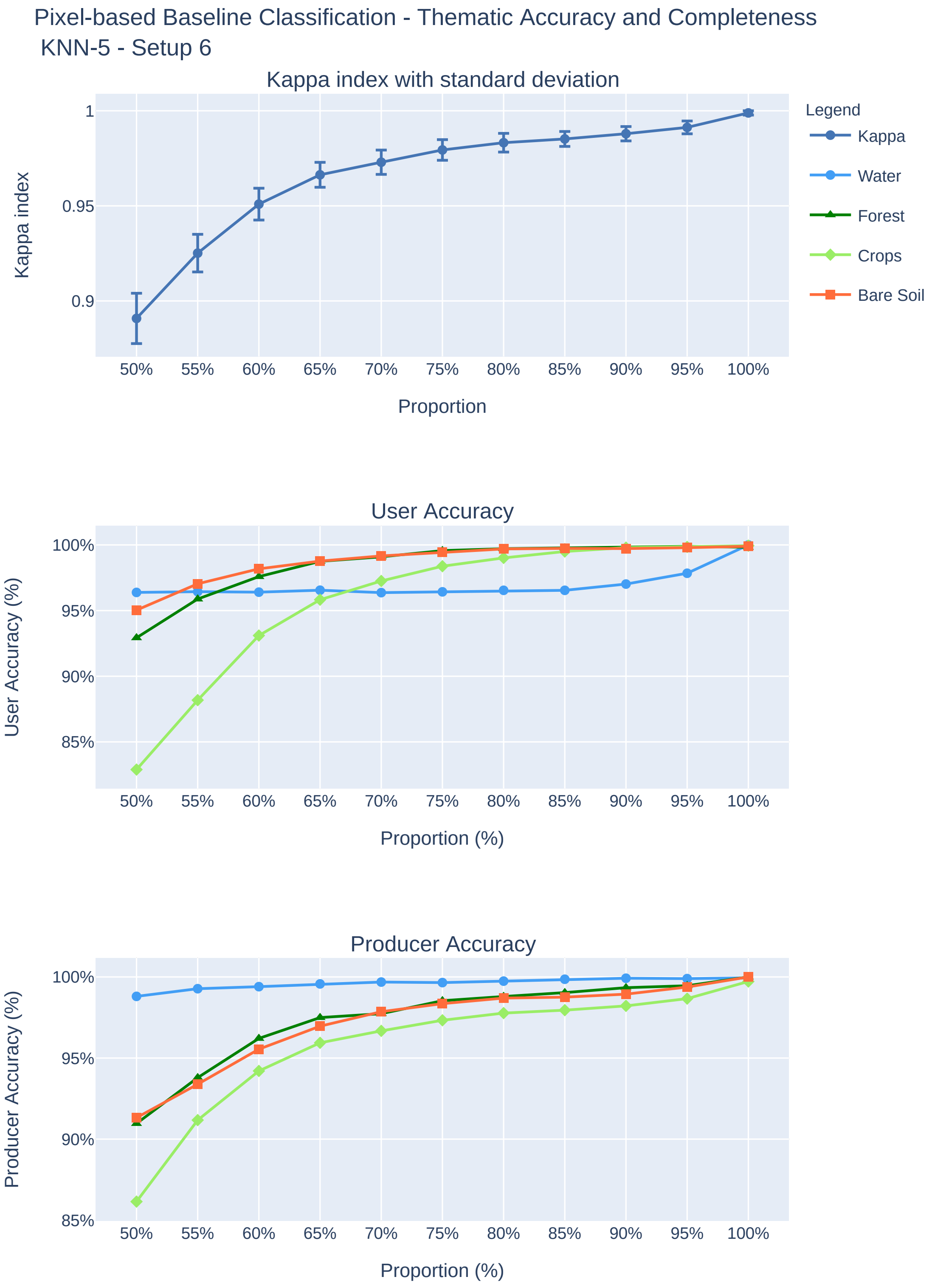
SOURCE: Author.

Figure B.11 - Pixel-based Baseline Classification - Accuracy Assessment for SVM-OAO for Setup 6.



SOURCE: Author.

Figure B.12 - Pixel-based Baseline Classification - Accuracy Assessment for KNN-5 for Setup 6.



SOURCE: Author.

B.1.2 Tabular results

The tabular results for Stage 1 Pixel-based Baseline Classification are presented in this appendix in two setups per page. Firstly the results regarding SVM-OAO classifier are presented followed by KNN-5 results.

The tables present the mean overall accuracy for the 100 Monte Carlo repetitions with the standard deviation. They also present the User Accuracy (UA) and Producer Accuracy (PA) in percentage for the four used classes.

These tables refer to Section 5.1.4, in page 73.

Table B.1 - Pixel-based Baseline Classification - Accuracy Assessment for SVM-OAO Setups 1 and 2.

Pixel-based Baseline Classification. Monte Carlo Simulation using SVM-OAO

Setup 1 Train: [prop, prop+0.05[

Test: [0.5, 1]

Interval		OA		*Kappa*		User Accuracy				Producer Accuracy			
Inf	**Sup**	**Mean**	**std**	**Mean**	**std**	**Water**	**Forest**	**Crops**	**Bare Soil**	**Water**	**Forest**	**Crops**	**Bare Soil**
50%	55%	0.7749	0.0075	0.6998	0.0100	96.53%	99.60%	13.95%	99.87%	98.41%	68.41%	96.30%	70.41%
55%	60%	0.8133	0.0098	0.7510	0.0131	96.54%	99.53%	29.44%	99.79%	98.55%	72.83%	96.61%	73.97%
60%	65%	0.8697	0.0094	0.8262	0.0125	96.29%	98.55%	53.25%	99.77%	98.98%	80.82%	95.02%	79.98%
65%	70%	0.9049	0.0091	0.8732	0.0122	96.31%	97.82%	70.58%	97.24%	99.11%	83.53%	91.83%	89.33%
70%	75%	0.9081	0.0089	0.8775	0.0119	96.10%	97.55%	74.02%	95.59%	99.29%	84.43%	89.68%	90.92%
75%	80%	0.9125	0.0094	0.8833	0.0125	95.86%	96.45%	77.98%	94.72%	99.34%	86.76%	87.76%	91.53%
80%	85%	0.9115	0.0093	0.8820	0.0124	95.41%	96.04%	79.45%	93.71%	99.45%	87.37%	86.16%	91.94%
85%	90%	0.9142	0.0092	0.8855	0.0123	94.83%	96.05%	80.94%	93.84%	99.42%	88.48%	86.07%	91.98%
90%	95%	0.9112	0.0099	0.8816	0.0132	93.93%	95.37%	82.30%	92.88%	99.58%	89.05%	83.95%	92.38%
95%	< 100%	0.9147	0.0085	0.8863	0.0114	94.75%	94.63%	84.06%	92.45%	99.61%	90.37%	83.93%	92.45%
100%	100%	0.9042	0.0111	0.8723	0.0148	92.23%	96.12%	82.12%	91.21%	99.71%	88.33%	81.78%	92.82%

Pixel-based Baseline Classification. Monte Carlo Simulation using SVM-OAO

Setup 2 Train: [prop, prop+0.05[

Test: [1]

Interval		OA		*Kappa*		User Accuracy				Producer Accuracy			
Inf	**Sup**	**Mean**	**std**	**Mean**	**std**	**Water**	**Forest**	**Crops**	**Bare Soil**	**Water**	**Forest**	**Crops**	**Bare Soil**
50%	55%	0.8550	0.0131	0.8067	0.0174	100.00%	100.00%	42.87%	99.13%	98.98%	82.49%	99.62%	73.59%
55%	60%	0.9400	0.0103	0.9200	0.0138	100.00%	100.00%	77.17%	98.83%	98.54%	92.69%	99.99%	87.11%
60%	65%	0.9909	0.0034	0.9879	0.0045	100.00%	100.00%	96.49%	99.87%	99.57%	99.48%	99.99%	97.38%
65%	70%	0.9979	0.0014	0.9971	0.0018	100.00%	100.00%	99.54%	99.60%	99.38%	99.84%	100.00%	99.92%
70%	75%	0.9978	0.0017	0.9971	0.0023	100.00%	100.00%	99.71%	99.42%	99.24%	99.91%	99.99%	100.00%
75%	80%	0.9990	0.0009	0.9987	0.0012	100.00%	100.00%	99.84%	99.78%	99.66%	100.00%	99.96%	100.00%
80%	85%	0.9982	0.0013	0.9976	0.0017	100.00%	100.00%	99.90%	99.39%	99.39%	100.00%	99.90%	100.00%
85%	90%	0.9985	0.0013	0.9980	0.0017	100.00%	100.00%	99.90%	99.51%	99.51%	100.00%	99.90%	100.00%
90%	95%	0.9987	0.0010	0.9983	0.0014	100.00%	100.00%	99.93%	99.55%	99.59%	100.00%	99.90%	100.00%
95%	< 100%	0.9992	0.0008	0.9990	0.0011	100.00%	100.00%	99.94%	99.75%	99.80%	100.00%	99.90%	100.00%
100%	100%	0.9996	0.0006	0.9994	0.0008	100.00%	100.00%	99.92%	99.90%	99.97%	99.95%	99.90%	100.00%

SOURCE: Author.

Table B.2 - Pixel-based Baseline Classification - Accuracy Assessment for KNN-5 Setups 1 and 2.

Pixel-based Baseline Classification. Monte Carlo Simulation using KNN-5
Setup 1 Train: [prop, prop+0.05[
Test: [0.5, 1]

Interval		OA		*Kappa*		User Accuracy				Producer Accuracy			
Inf	**Sup**	**Mean**	**std**	**Mean**	**std**	**Water**	**Forest**	**Crops**	**Bare Soil**	**Water**	**Forest**	**Crops**	**Bare Soil**
50%	55%	0.8099	0.0128	0.7466	0.0170	99.82%	93.38%	48.36%	82.41%	96.92%	75.78%	70.01%	78.70%
55%	60%	0.8820	0.0106	0.8426	0.0141	98.87%	92.77%	67.79%	93.36%	98.65%	83.55%	84.71%	85.89%
60%	65%	0.9124	0.0113	0.8832	0.0151	97.56%	91.80%	81.62%	93.99%	99.25%	90.52%	86.07%	89.13%
65%	70%	0.9128	0.0110	0.8837	0.0147	96.34%	90.10%	85.74%	92.94%	99.44%	92.05%	84.03%	90.06%
70%	75%	0.9093	0.0114	0.8791	0.0152	95.24%	89.60%	87.51%	91.37%	99.44%	90.82%	82.66%	91.65%
75%	80%	0.9070	0.0119	0.8760	0.0159	95.22%	88.09%	88.74%	90.75%	99.52%	92.17%	81.15%	91.30%
80%	85%	0.8998	0.0122	0.8663	0.0163	94.91%	86.52%	89.04%	89.43%	99.37%	93.08%	79.41%	89.96%
85%	90%	0.8782	0.0143	0.8376	0.0191	90.56%	83.89%	90.82%	86.01%	99.53%	95.24%	75.73%	85.15%
90%	95%	0.8781	0.0137	0.8375	0.0182	90.46%	83.75%	90.03%	87.00%	99.60%	94.03%	76.02%	85.56%
95%	< 100%	0.8537	0.0173	0.8049	0.0231	86.19%	80.42%	92.34%	82.53%	99.68%	96.80%	72.14%	80.55%
100%	100%	0.8140	0.0181	0.7520	0.0242	75.94%	75.96%	94.52%	79.19%	99.92%	98.18%	68.39%	73.03%

Pixel-based Baseline Classification. Monte Carlo Simulation using KNN-5
Setup 2 Train: [prop, prop+0.05[
Test: [1]

Interval		OA		*Kappa*		User Accuracy				Producer Accuracy			
Inf	**Sup**	**Mean**	**std**	**Mean**	**std**	**Water**	**Forest**	**Crops**	**Bare Soil**	**Water**	**Forest**	**Crops**	**Bare Soil**
50%	55%	0.9150	0.0158	0.8866	0.0211	100.00%	99.08%	75.26%	91.64%	97.90%	86.17%	91.38%	91.17%
55%	60%	0.9665	0.0110	0.9553	0.0146	100.00%	99.97%	87.85%	98.78%	99.12%	90.76%	99.67%	97.96%
60%	65%	0.9945	0.0036	0.9927	0.0048	100.00%	100.00%	98.26%	99.55%	99.70%	98.92%	99.92%	99.27%
65%	70%	0.9987	0.0013	0.9982	0.0017	100.00%	100.00%	99.56%	99.92%	99.88%	99.92%	100.00%	99.67%
70%	75%	0.9990	0.0008	0.9986	0.0011	100.00%	100.00%	99.78%	99.81%	99.85%	99.85%	99.96%	99.92%
75%	80%	0.9995	0.0006	0.9993	0.0008	100.00%	100.00%	99.93%	99.87%	99.97%	99.94%	99.91%	99.98%
80%	85%	0.9995	0.0007	0.9993	0.0009	100.00%	100.00%	99.94%	99.85%	99.88%	99.97%	99.97%	99.97%
85%	90%	0.9996	0.0005	0.9995	0.0007	100.00%	100.00%	99.87%	99.97%	99.98%	99.95%	99.99%	99.92%
90%	95%	0.9998	0.0005	0.9997	0.0006	100.00%	100.00%	99.94%	99.97%	99.98%	99.99%	100.00%	99.94%
95%	< 100%	0.9998	0.0004	0.9998	0.0005	100.00%	100.00%	99.96%	99.98%	100.00%	99.97%	99.98%	99.99%
100%	100%	0.9996	0.0005	0.9995	0.0007	100.00%	99.98%	99.89%	99.97%	100.00%	100.00%	99.95%	99.89%

SOURCE: Author.

Table B.3 - Pixel-based Baseline Classification - Accuracy Assessment for SVM-OAO Setups 3 and 4.

Pixel-based Baseline Classification. Monte Carlo Simulation using SVM-OAO

Setup 3 Train: [prop, 1]

Test: [0.5,1]

Threshold	OA		*Kappa*		User Accuracy				Producer Accuracy			
	Mean	std	Mean	std	Water	Forest	Crops	Bare Soil	Water	Forest	Crops	Bare Soil
50%	0.8984	0.0083	0.8645	0.0111	96.46%	97.81%	67.60%	97.47%	98.61%	82.40%	92.28%	88.43%
55%	0.9037	0.0096	0.8716	0.0128	96.39%	97.56%	70.93%	96.62%	98.84%	83.68%	91.03%	89.48%
60%	0.9084	0.0091	0.8778	0.0122	96.30%	97.24%	73.82%	95.98%	99.07%	84.82%	90.05%	90.41%
65%	0.9095	0.0088	0.8793	0.0118	96.24%	96.96%	75.44%	95.17%	99.18%	85.34%	88.92%	91.06%
70%	0.9108	0.0089	0.8810	0.0119	96.08%	96.61%	77.09%	94.52%	99.32%	86.19%	87.87%	91.38%
75%	0.9110	0.0092	0.8813	0.0123	95.64%	96.36%	78.35%	94.03%	99.40%	86.77%	86.91%	91.67%
80%	0.9119	0.0097	0.8825	0.0130	95.20%	95.94%	80.00%	93.60%	99.50%	87.80%	85.91%	91.84%
85%	0.9114	0.0099	0.8819	0.0132	94.39%	95.86%	81.13%	93.19%	99.64%	88.29%	84.85%	92.20%
90%	0.9096	0.0105	0.8795	0.0139	93.50%	95.47%	82.11%	92.78%	99.66%	88.85%	83.53%	92.36%
95%	0.9109	0.0096	0.8812	0.0128	93.73%	95.16%	83.00%	92.47%	99.66%	89.51%	83.39%	92.39%
100%	0.9042	0.0111	0.8723	0.0148	92.23%	96.12%	82.12%	91.21%	99.71%	88.33%	81.78%	92.82%

Pixel-based Baseline Classification. Monte Carlo Simulation using SVM-OAO

Setup 4 Train: [prop, 1]

Test: [1]

Threshold	OA		*Kappa*		User Accuracy				Producer Accuracy			
	Mean	std	Mean	std	Water	Forest	Crops	Bare Soil	Water	Forest	Crops	Bare Soil
50%	0.9973	0.0016	0.9964	0.0021	100.00%	100.00%	99.25%	99.68%	99.36%	99.69%	100.00%	99.88%
55%	0.9980	0.0015	0.9973	0.0020	100.00%	100.00%	99.50%	99.69%	99.42%	99.84%	99.98%	99.96%
60%	0.9983	0.0013	0.9978	0.0017	100.00%	100.00%	99.66%	99.67%	99.47%	99.90%	99.97%	100.00%
65%	0.9984	0.0014	0.9979	0.0019	100.00%	100.00%	99.73%	99.63%	99.51%	99.92%	99.93%	100.00%
70%	0.9985	0.0013	0.9980	0.0018	100.00%	100.00%	99.78%	99.63%	99.55%	99.93%	99.93%	100.00%
75%	0.9987	0.0011	0.9983	0.0015	100.00%	100.00%	99.84%	99.65%	99.62%	99.98%	99.90%	100.00%
80%	0.9990	0.0011	0.9987	0.0014	100.00%	100.00%	99.91%	99.68%	99.70%	100.00%	99.90%	100.00%
85%	0.9991	0.0009	0.9988	0.0013	100.00%	100.00%	99.93%	99.70%	99.73%	100.00%	99.90%	100.00%
90%	0.9993	0.0008	0.9991	0.0010	100.00%	100.00%	99.94%	99.79%	99.84%	100.00%	99.89%	100.00%
95%	0.9995	0.0006	0.9993	0.0008	100.00%	100.00%	99.95%	99.85%	99.91%	100.00%	99.90%	100.00%
100%	0.9996	0.0006	0.9994	0.0008	100.00%	100.00%	99.92%	99.90%	99.97%	99.95%	99.90%	100.00%

SOURCE: Author.

Table B.4 - Pixel-based Baseline Classification - Accuracy Assessment for KNN-5 Setups 3 and 4.

Pixel-based Baseline Classification. Monte Carlo Simulation using KNN-5
Setup 3 Train: [prop, 1]
Test: [0.5,1]

Threshold	OA		*Kappa*		User Accuracy				Producer Accuracy			
	Mean	**std**	**Mean**	**std**	**Water**	**Forest**	**Crops**	**Bare Soil**	**Water**	**Forest**	**Crops**	**Bare Soil**
50%	0.9110	0.0108	0.8813	0.0144	99.27%	90.60%	82.62%	91.90%	98.48%	91.78%	83.95%	90.04%
55%	0.9126	0.0108	0.8834	0.0144	98.42%	90.26%	84.56%	91.78%	98.97%	92.34%	83.41%	90.49%
60%	0.9106	0.0116	0.8808	0.0155	97.74%	89.28%	86.02%	91.21%	99.18%	92.54%	82.27%	90.82%
65%	0.9087	0.0114	0.8783	0.0152	97.36%	88.72%	87.20%	90.21%	99.22%	92.83%	81.20%	91.20%
70%	0.9036	0.0120	0.8714	0.0161	96.48%	87.68%	87.92%	89.34%	99.34%	92.72%	79.87%	90.95%
75%	0.8977	0.0121	0.8636	0.0161	95.63%	86.40%	88.59%	88.45%	99.43%	93.47%	78.52%	89.73%
80%	0.8872	0.0137	0.8496	0.0183	93.08%	85.04%	89.49%	87.26%	99.51%	94.59%	77.03%	86.94%
85%	0.8725	0.0161	0.8301	0.0215	89.60%	83.31%	90.43%	85.67%	99.71%	95.17%	75.11%	83.83%
90%	0.8592	0.0170	0.8122	0.0226	85.58%	82.49%	91.02%	84.57%	99.79%	95.67%	74.05%	80.46%
95%	0.8395	0.0180	0.7860	0.0240	81.42%	79.98%	92.38%	82.01%	99.87%	96.66%	71.56%	76.91%
100%	0.8140	0.0181	0.7520	0.0242	75.94%	75.96%	94.52%	79.19%	99.92%	98.18%	68.39%	73.03%

Pixel-based Baseline Classification. Monte Carlo Simulation using KNN-5
Setup 4 Train: [prop, 1]
Test: [1]

Threshold	OA		*Kappa*		User Accuracy				Producer Accuracy			
	Mean	**std**	**Mean**	**std**	**Water**	**Forest**	**Crops**	**Bare Soil**	**Water**	**Forest**	**Crops**	**Bare Soil**
50%	0.6851	0.0257	0.5802	0.0343	79.70%	37.18%	57.86%	99.30%	67.15%	43.27%	67.82%	90.21%
55%	0.7691	0.0174	0.6922	0.0232	91.04%	36.22%	80.71%	99.68%	74.25%	61.16%	69.25%	98.10%
60%	0.8312	0.0114	0.7750	0.0152	95.70%	49.68%	87.44%	99.68%	81.39%	79.58%	73.87%	98.08%
65%	0.8527	0.0106	0.8036	0.0141	98.13%	54.24%	89.04%	99.66%	84.32%	82.88%	75.63%	99.23%
70%	0.8648	0.0102	0.8197	0.0136	99.12%	60.67%	86.47%	99.66%	86.84%	83.29%	76.72%	99.35%
75%	0.8888	0.0088	0.8518	0.0117	98.75%	64.85%	92.13%	99.80%	88.62%	90.40%	79.21%	99.28%
80%	0.8976	0.0088	0.8634	0.0117	99.09%	70.72%	89.65%	99.57%	91.50%	87.43%	80.89%	99.58%
85%	0.9131	0.0085	0.8841	0.0114	99.22%	75.19%	91.21%	99.61%	91.98%	90.21%	83.92%	99.51%
90%	0.9164	0.0074	0.8885	0.0099	99.14%	77.77%	90.27%	99.36%	93.62%	89.12%	84.27%	99.63%
95%	0.9186	0.0076	0.8914	0.0101	99.20%	77.87%	90.68%	99.67%	94.94%	89.66%	83.58%	99.50%
100%	0.9486	0.0070	0.9314	0.0093	98.48%	89.18%	92.22%	99.55%	98.53%	91.29%	90.15%	99.49%

SOURCE: Author.

Table B.5 - Pixel-based Baseline Classification - Accuracy Assessment for SVM-OAO Setups 5 and 6.

Pixel-based Baseline Classification. Monte Carlo Simulation using SVM-OAO
Setup 5 Train: [prop, prop+0.05[
Test: [prop, 1]

Interval		OA		*Kappa*		User Accuracy				Producer Accuracy			
Inf	Sup	Mean	std	Mean	std	Water	Forest	Crops	Bare Soil	Water	Forest	Crops	Bare Soil
50%	55%	0.7743	0.0066	0.6991	0.0087	96.54%	99.59%	13.72%	99.88%	98.37%	68.42%	96.56%	70.28%
55%	60%	0.8189	0.0089	0.7585	0.0119	96.60%	99.84%	31.32%	99.79%	98.91%	73.44%	97.40%	74.34%
60%	65%	0.8939	0.0088	0.8585	0.0117	96.34%	99.70%	61.55%	99.96%	99.31%	84.76%	97.61%	81.73%
65%	70%	0.9496	0.0064	0.9328	0.0086	96.55%	99.81%	83.96%	99.53%	99.46%	90.18%	97.75%	93.58%
70%	75%	0.9641	0.0055	0.9522	0.0073	96.39%	99.86%	90.06%	99.34%	99.57%	92.56%	97.74%	96.29%
75%	80%	0.9766	0.0045	0.9688	0.0060	96.43%	99.87%	94.76%	99.59%	99.66%	95.79%	97.89%	97.46%
80%	85%	0.9824	0.0043	0.9765	0.0057	96.51%	99.89%	96.89%	99.68%	99.72%	97.21%	98.07%	98.04%
85%	90%	0.9869	0.0033	0.9826	0.0044	96.51%	99.94%	98.65%	99.67%	99.79%	98.32%	98.31%	98.39%
90%	95%	0.9899	0.0028	0.9865	0.0038	97.02%	100.00%	99.36%	99.59%	99.81%	99.01%	98.37%	98.80%
95%	< 100%	0.9934	0.0023	0.9912	0.0031	97.86%	100.00%	99.81%	99.68%	99.84%	99.38%	98.78%	99.37%
100%	100%	0.9994	0.0007	0.9992	0.0009	100.00%	100.00%	99.94%	99.83%	99.94%	100.00%	99.83%	100.00%

Pixel-based Baseline Classification. Monte Carlo Simulation using SVM-OAO
Setup 6 Train: [prop, 1]
Test: [prop, 1]

Threshold	OA		*Kappa*		User Accuracy				Producer Accuracy			
	Mean	std	Mean	std	Water	Forest	Crops	Bare Soil	Water	Forest	Crops	Bare Soil
50%	0.8978	0.0097	0.8638	0.0130	96.46%	97.84%	67.40%	97.43%	98.62%	82.18%	92.27%	88.50%
55%	0.9209	0.0092	0.8945	0.0122	96.40%	98.64%	74.98%	98.35%	99.21%	85.47%	94.63%	90.89%
60%	0.9432	0.0081	0.9243	0.0108	96.41%	99.26%	82.87%	98.74%	99.41%	89.36%	96.08%	93.43%
65%	0.9589	0.0062	0.9452	0.0083	96.54%	99.64%	88.47%	98.91%	99.56%	92.21%	96.92%	95.38%
70%	0.9684	0.0054	0.9579	0.0073	96.36%	99.72%	92.04%	99.23%	99.69%	94.02%	97.34%	96.61%
75%	0.9768	0.0044	0.9691	0.0059	96.42%	99.88%	94.95%	99.49%	99.69%	95.89%	97.78%	97.52%
80%	0.9830	0.0042	0.9773	0.0056	96.44%	99.93%	97.15%	99.67%	99.81%	97.45%	98.03%	97.98%
85%	0.9870	0.0033	0.9827	0.0045	96.50%	99.92%	98.75%	99.64%	99.88%	98.34%	98.16%	98.48%
90%	0.9900	0.0029	0.9867	0.0039	97.01%	99.99%	99.36%	99.64%	99.94%	98.99%	98.34%	98.76%
95%	0.9933	0.0023	0.9911	0.0030	97.85%	100.00%	99.74%	99.74%	99.90%	99.32%	98.77%	99.36%
100%	0.9994	0.0007	0.9992	0.0009	100.00%	100.00%	99.94%	99.83%	99.94%	100.00%	99.83%	100.00%

SOURCE: Author.

Table B.6 - Pixel-based Baseline Classification - Accuracy Assessment for KNN-5 Setups 5 and 6.

Pixel-based Baseline Classification. Monte Carlo Simulation using KNN-5
Setup 5 Train: [prop, prop+0.05[
Test: [prop,1]

Interval		OA		*Kappa*		User Accuracy				Producer Accuracy			
Inf	**Sup**	**Mean**	**std**	**Mean**	**std**	**Water**	**Forest**	**Crops**	**Bare Soil**	**Water**	**Forest**	**Crops**	**Bare Soil**
50%	55%	0.8007	0.0154	0.7343	0.0205	96.49%	90.76%	44.95%	88.08%	97.47%	76.05%	68.31%	76.03%
55%	60%	0.9183	0.0092	0.8910	0.0122	96.55%	96.48%	77.16%	97.12%	98.57%	90.01%	91.76%	87.67%
60%	65%	0.9579	0.0068	0.9439	0.0091	96.40%	98.29%	89.93%	98.56%	99.15%	95.66%	95.15%	93.40%
65%	70%	0.9770	0.0044	0.9694	0.0058	96.55%	99.24%	95.94%	99.09%	99.36%	97.73%	96.98%	96.80%
70%	75%	0.9827	0.0039	0.9769	0.0052	96.33%	99.45%	97.75%	99.54%	99.76%	98.17%	97.32%	97.89%
75%	80%	0.9865	0.0034	0.9820	0.0045	96.42%	99.70%	98.86%	99.63%	99.79%	98.71%	97.63%	98.53%
80%	85%	0.9879	0.0040	0.9838	0.0053	96.53%	99.63%	99.36%	99.63%	99.64%	98.99%	97.75%	98.80%
85%	90%	0.9891	0.0030	0.9854	0.0040	96.68%	99.75%	99.39%	99.81%	99.79%	99.11%	98.02%	98.75%
90%	95%	0.9909	0.0027	0.9878	0.0037	97.02%	99.80%	99.73%	99.80%	99.86%	99.41%	98.15%	98.97%
95%	< 100%	0.9939	0.0023	0.9919	0.0031	97.85%	99.96%	99.89%	99.87%	99.90%	99.49%	98.83%	99.36%
100%	100%	0.9992	0.0009	0.9989	0.0011	100.00%	99.82%	99.94%	99.90%	99.94%	100.00%	99.72%	100.00%

Pixel-based Baseline Classification. Monte Carlo Simulation using KNN-5
Setup 6 Train: [prop, 1]
Test: [prop, 1]

Threshold	OA		*Kappa*		User Accuracy				Producer Accuracy			
	Mean	**std**	**Mean**	**std**	**Water**	**Forest**	**Crops**	**Bare Soil**	**Water**	**Forest**	**Crops**	**Bare Soil**
50%	0.9181	0.0099	0.8908	0.0132	96.38%	92.93%	82.89%	95.04%	98.80%	90.96%	86.15%	91.32%
55%	0.9439	0.0074	0.9251	0.0099	96.44%	95.88%	88.18%	97.04%	99.28%	93.79%	91.17%	93.39%
60%	0.9632	0.0063	0.9509	0.0084	96.40%	97.59%	93.10%	98.18%	99.40%	96.21%	94.20%	95.55%
65%	0.9748	0.0049	0.9663	0.0066	96.55%	98.75%	95.83%	98.77%	99.57%	97.50%	95.94%	96.96%
70%	0.9797	0.0048	0.9729	0.0064	96.37%	99.10%	97.26%	99.16%	99.69%	97.74%	96.68%	97.86%
75%	0.9846	0.0041	0.9794	0.0054	96.43%	99.57%	98.38%	99.44%	99.65%	98.54%	97.33%	98.35%
80%	0.9874	0.0037	0.9832	0.0049	96.54%	99.71%	99.01%	99.70%	99.74%	98.80%	97.77%	98.70%
85%	0.9889	0.0029	0.9852	0.0039	96.54%	99.78%	99.50%	99.74%	99.86%	99.04%	97.95%	98.76%
90%	0.9910	0.0028	0.9879	0.0038	97.02%	99.84%	99.80%	99.72%	99.92%	99.34%	98.22%	98.94%
95%	0.9935	0.0025	0.9913	0.0034	97.84%	99.90%	99.85%	99.79%	99.90%	99.46%	98.66%	99.38%
100%	0.9992	0.0009	0.9989	0.0011	100.00%	99.82%	99.94%	99.90%	99.94%	100.00%	99.72%	100.00%

SOURCE: Author.

B.1.3 Confusion matrices

The confusion matrices regarding pixel-based baseline classification are presented in this appendix. Firstly the results of the SVM-OAO classifier is presented followed by KNN-5.

Table B.7 - Pixel-based Baseline Classification - Confusion Matrix for SVM-OAO Setup 1.

Pixel-based Baseline Classification. Monte Carlo Simulation using SVM-OAO
Setup 1 Train: [prop, prop + 5%[
Test: [50% 100%]

50-55%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**241.32**	3.73	1	3.95	250
Grass	0.77	**249**	0.22	0.01	250
Bare Soil	2.91	111.24	**34.88**	100.97	250
Water	0.21	0	0.12	**249.67**	250
Total	245.21	363.97	36.22	354.6	1000

55-60%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**241.35**	3.43	1.74	3.48	250
Grass	0.32	**248.83**	0.84	0.01	250
Bare Soil	2.69	89.42	**73.6**	84.29	250
Water	0.53	0	0	**249.47**	250
Total	244.89	341.68	76.18	337.25	1000

60-65%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**240.73**	2.91	2.99	3.37	250
Grass	0.15	**246.37**	3.48	0	250
Bare Soil	2.28	55.54	**133.13**	59.05	250
Water	0.06	0	0.51	**249.43**	250
Total	243.22	304.82	140.11	311.85	1000

65-70%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**240.78**	2.86	3.63	2.73	250
Grass	0.1	**244.55**	5.35	0	250
Bare Soil	1.9	45.35	**176.44**	26.31	250
Water	0.17	0	6.72	**243.11**	250
Total	242.95	292.76	192.14	272.15	1000

70-75%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**240.24**	2.81	4.32	2.63	250
Grass	0.02	**243.88**	6.1	0	250
Bare Soil	1.56	42.17	**185.04**	21.23	250
Water	0.15	0	10.88	**238.97**	250
Total	241.97	288.86	206.34	262.83	1000

75-80%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**239.65**	2.64	5.22	2.49	250
Grass	0.01	**241.12**	8.87	0	250
Bare Soil	1.47	34.16	**194.94**	19.43	250
Water	0.11	0	13.09	**236.8**	250
Total	241.24	277.92	222.12	258.72	1000

80-85%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**238.52**	2.55	6.48	2.45	250
Grass	0.03	**240.09**	9.88	0	250
Bare Soil	1.12	32.17	**198.62**	18.09	250
Water	0.17	0	15.55	**234.28**	250
Total	239.84	274.81	230.53	254.82	1000

85-90%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**237.08**	2.54	8	2.38	250
Grass	0.39	**240.13**	9.48	0	250
Bare Soil	0.87	28.71	**202.34**	18.08	250
Water	0.13	0	15.27	**234.6**	250
Total	238.47	271.38	235.09	255.06	1000

90-95%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**234.83**	2.53	10.11	2.53	250
Grass	0.01	**238.42**	11.57	0	250
Bare Soil	0.84	26.78	**205.75**	16.63	250
Water	0.14	0	17.65	**232.21**	250
Total	235.82	267.73	245.08	251.37	1000

95-100%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**236.88**	2.44	8.04	2.64	250
Grass	0.03	**236.58**	13.39	0	250
Bare Soil	0.84	22.78	**210.15**	16.23	250
Water	0.05	0	18.82	**231.13**	250
Total	237.8	261.8	250.4	250	1000

100%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**230.58**	2.83	14.06	2.53	250
Grass	0	**240.3**	9.7	0	250
Bare Soil	0.68	28.92	**205.29**	15.11	250
Water	0	0	21.97	**228.03**	250
Total	231.26	272.05	251.02	245.67	1000

SOURCE: Author.

Table B.8 - Pixel-based Baseline Classification - Confusion Matrix for KNN-5 Setup 1.

Pixel-based Baseline Classification. Monte Carlo Simulation using KNN-5

Setup 1 Train: [prop, prop + 5%[

Test: [50%, 100%]

50-55% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**249.54**	0.2	0.09	0.17	250
Grass	4.02	**233.45**	12.3	0.23	250
Bare Soil	0	73.75	**120.9**	55.35	250
Water	3.9	0.68	39.39	**206.03**	250
Total	257.46	308.08	172.68	261.78	1000

55-60% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**247.18**	2.22	0.08	0.52	250
Grass	0.85	**231.93**	17	0.22	250
Bare Soil	0	42.92	**169.47**	37.61	250
Water	2.54	0.54	13.52	**233.4**	250
Total	250.57	277.61	200.07	271.75	1000

60-65% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**243.9**	4.19	0.08	1.83	250
Grass	0.62	**229.51**	19.67	0.2	250
Bare Soil	0	19.33	**204.05**	26.62	250
Water	1.23	0.52	13.28	**234.97**	250
Total	245.75	253.55	237.08	263.62	1000

65-70% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**240.85**	6.64	0.08	2.43	250
Grass	0.49	**225.25**	24.04	0.22	250
Bare Soil	0.04	12.63	**214.34**	22.99	250
Water	0.82	0.19	16.63	**232.36**	250
Total	242.2	244.71	255.09	258	1000

70-75% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**238.1**	9.53	0.06	2.31	250
Grass	0.3	**224.01**	25.44	0.25	250
Bare Soil	0	12.98	**218.77**	18.25	250
Water	1.03	0.14	20.4	**228.43**	250
Total	239.43	246.66	264.67	249.24	1000

75-80% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**238.04**	7.49	0.16	4.31	250
Grass	0.39	**220.22**	29.18	0.21	250
Bare Soil	0	11.05	**221.85**	17.1	250
Water	0.76	0.17	22.2	**226.87**	250
Total	239.19	238.93	273.39	248.49	1000

80-85% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**237.28**	6.02	0.08	6.62	250
Grass	0.33	**216.3**	32.44	0.93	250
Bare Soil	0	10	**222.59**	17.41	250
Water	1.17	0.06	25.19	**223.58**	250
Total	238.78	232.38	280.3	248.54	1000

85-90% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**226.39**	1.91	0.13	21.57	250
Grass	0.16	**209.73**	38.75	1.36	250
Bare Soil	0	8.39	**227.04**	14.57	250
Water	0.92	0.18	33.88	**215.02**	250
Total	227.47	220.21	299.8	252.52	1000

90-95% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**226.14**	3.75	0.11	20	250
Grass	0.19	**209.37**	39.36	1.08	250
Bare Soil	0	9.31	**225.07**	15.62	250
Water	0.72	0.24	31.53	**217.51**	250
Total	227.05	222.67	296.07	254.21	1000

95-100% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**215.48**	0.16	0.19	34.17	250
Grass	0.12	**201.05**	45.87	2.96	250
Bare Soil	0	6.47	**230.85**	12.68	250
Water	0.57	0.02	43.09	**206.32**	250
Total	216.17	207.7	320	256.13	1000

100% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**189.85**	0.08	0.21	59.86	250
Grass	0.08	**189.89**	57.06	2.97	250
Bare Soil	0	3.44	**236.29**	10.27	250
Water	0.07	0	51.96	**197.97**	250
Total	190	193.41	345.52	271.07	1000

SOURCE: Author.

Table B.9 - Pixel-based Baseline Classification - Confusion Matrix for SVM-OAO Setup 2.

Pixel-based Baseline Classification. Monte Carlo Simulation using SVM-OAO
Setup 2 Train: [prop, prop + 5%[
Test: [100%]

50-55% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0.82	53.06	**107.18**	88.94	250
Water	1.76	0	0.41	**247.83**	250
Total	252.58	303.06	107.59	336.77	1000

55-60% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0.78	19.73	**192.93**	36.56	250
Water	2.92	0	0.01	**247.07**	250
Total	253.7	269.73	192.94	283.63	1000

60-65% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0.76	1.3	**241.23**	6.71	250
Water	0.31	0	0.02	**249.67**	250
Total	251.07	251.3	241.25	256.38	1000

65-70% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0.56	0.39	**248.86**	0.19	250
Water	1.01	0	0	**248.99**	250
Total	251.57	250.39	248.86	249.18	1000

70-75% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0.5	0.22	**249.28**	0	250
Water	1.41	0	0.03	**248.56**	250
Total	251.91	250.22	249.31	248.56	1000

75-80% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0.4	0.01	**249.59**	0	250
Water	0.45	0	0.1	**249.45**	250
Total	250.85	250.01	249.69	249.45	1000

80-85% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0.26	0	**249.74**	0	250
Water	1.27	0	0.26	**248.47**	250
Total	251.53	250	250	248.47	1000

85-90% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0.25	0	**249.75**	0	250
Water	0.98	0	0.25	**248.77**	250
Total	251.23	250	250	248.77	1000

90-95% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0.18	0	**249.82**	0	250
Water	0.86	0	0.26	**248.88**	250
Total	251.04	250	250.08	248.88	1000

95-100% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0.14	0	**249.86**	0	250
Water	0.36	0	0.26	**249.38**	250
Total	250.5	250	250.12	249.38	1000

100% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0.07	0.12	**249.81**	0	250
Water	0	0	0.26	**249.74**	250
Total	250.07	250.12	250.07	249.74	1000

SOURCE: Author.

Table B.10 - Pixel-based Baseline Classification - Confusion Matrix for KNN-5 Setup 2.

Pixel-based Baseline Classification. Monte Carlo Simulation using KNN-5

Setup 2 Train: [prop, prop + 5%[

Test: [100%]

50-55%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**250**	0	0	0	250
Grass	0	**247.71**	2.29	0	250
Bare Soil	0	39.66	**188.15**	22.19	250
Water	5.35	0.1	15.46	**229.09**	250
Total	255.35	287.47	205.9	251.28	1000

55-60%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**250**	0	0	0	250
Grass	0	**249.93**	0.07	0	250
Bare Soil	0	25.25	**219.62**	5.13	250
Water	2.21	0.19	0.66	**246.94**	250
Total	252.21	275.37	220.35	252.07	1000

60-65%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0	2.52	**245.66**	1.82	250
Water	0.74	0.2	0.19	**248.87**	250
Total	250.74	252.72	245.85	250.69	1000

65-70%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0.09	0.2	**248.89**	0.82	250
Water	0.2	0	0.01	**249.79**	250
Total	250.29	250.2	248.9	250.61	1000

70-75%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0	0.37	**249.44**	0.19	250
Water	0.38	0	0.1	**249.52**	250
Total	250.38	250.37	249.54	249.71	1000

75-80%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0	0.12	**249.82**	0.06	250
Water	0.07	0.03	0.22	**249.68**	250
Total	250.07	250.15	250.04	249.74	1000

80-85%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0	0.08	**249.84**	0.08	250
Water	0.3	0	0.07	**249.63**	250
Total	250.3	250.08	249.91	249.71	1000

85-90%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0	0.13	**249.68**	0.19	250
Water	0.04	0	0.03	**249.93**	250
Total	250.04	250.13	249.71	250.12	1000

90-95%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0	0.01	**249.85**	0.14	250
Water	0.06	0.01	0	**249.93**	250
Total	250.06	250.02	249.85	250.07	1000

95-100%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0	0.07	**249.9**	0.03	250
Water	0	0	0.06	**249.94**	250
Total	250	250.07	249.96	249.97	1000

100%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**250**	0	0	0	250
Grass	0	**249.95**	0.05	0	250
Bare Soil	0	0	**249.72**	0.28	250
Water	0	0	0.07	**249.93**	250
Total	250	249.95	249.84	250.21	1000

SOURCE: Author.

Table B.11 - Pixel-based Baseline Classification - Confusion Matrix for SVM-OAO Setup 3.

Pixel-based Baseline Classification. Monte Carlo Simulation using SVM-OAO
Setup 3 Train: [prop, 100%]
Test: [50% 100%]

50% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**241.15**	2.82	3.24	2.79	250
Grass	0.69	**244.53**	4.78	0	250
Bare Soil	2.5	49.41	**169**	29.09	250
Water	0.2	0	6.12	**243.68**	250
Total	244.54	296.76	183.14	275.56	1000

55% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**240.97**	2.8	3.45	2.78	250
Grass	0.39	**243.89**	5.72	0	250
Bare Soil	2.29	44.77	**177.32**	25.62	250
Water	0.14	0	8.31	**241.55**	250
Total	243.79	291.46	194.8	269.95	1000

60% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**240.74**	2.73	3.75	2.78	250
Grass	0.2	**243.1**	6.7	0	250
Bare Soil	1.97	40.79	**184.56**	22.68	250
Water	0.09	0	9.95	**239.96**	250
Total	243	286.62	204.96	265.42	1000

65% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**240.61**	2.73	4.07	2.59	250
Grass	0.12	**242.39**	7.49	0	250
Bare Soil	1.73	38.91	**188.59**	20.77	250
Water	0.13	0	11.95	**237.92**	250
Total	242.59	284.03	212.1	261.28	1000

70% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**240.2**	2.66	4.58	2.56	250
Grass	0.09	**241.52**	8.39	0	250
Bare Soil	1.5	36.04	**192.73**	19.73	250
Water	0.06	0	13.63	**236.31**	250
Total	241.85	280.22	219.33	258.6	1000

75% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**239.11**	2.63	5.67	2.59	250
Grass	0.11	**240.91**	8.98	0	250
Bare Soil	1.27	34.09	**195.88**	18.76	250
Water	0.07	0	14.85	**235.08**	250
Total	240.56	277.63	225.38	256.43	1000

80% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**238.01**	2.55	6.82	2.62	250
Grass	0.09	**239.86**	10.05	0	250
Bare Soil	1.05	30.79	**200**	18.16	250
Water	0.06	0	15.94	**234**	250
Total	239.21	273.2	232.81	254.78	1000

85% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**235.97**	2.55	8.91	2.57	250
Grass	0.02	**239.66**	10.32	0	250
Bare Soil	0.8	29.23	**202.83**	17.14	250
Water	0.03	0	16.99	**232.98**	250
Total	236.82	271.44	239.05	252.69	1000

90% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**233.74**	2.57	11.12	2.57	250
Grass	0.03	**238.67**	11.3	0	250
Bare Soil	0.74	27.37	**205.27**	16.62	250
Water	0.02	0	18.04	**231.94**	250
Total	234.53	268.61	245.73	251.13	1000

95% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**234.33**	2.62	10.44	2.61	250
Grass	0.01	**237.91**	12.08	0	250
Bare Soil	0.78	25.27	**207.51**	16.44	250
Water	0	0	18.82	**231.18**	250
Total	235.12	265.8	248.85	250.23	1000

100% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**230.58**	2.83	14.06	2.53	250
Grass	0	**240.3**	9.7	0	250
Bare Soil	0.68	28.92	**205.29**	15.11	250
Water	0	0	21.97	**228.03**	250
Total	231.26	272.05	251.02	245.67	1000

SOURCE: Author.

Table B.12 - Pixel-based Baseline Classification - Confusion Matrix for KNN-5 Setup 3.

Pixel-based Baseline Classification. Monte Carlo Simulation using KNN-5
Setup 3 Train: [prop, 100%]
Test: [50%, 100%]

50% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**248.17**	0.72	0.12	0.99	250
Grass	1.74	**226.5**	21.56	0.2	250
Bare Soil	0.01	19.21	**206.56**	24.22	250
Water	2.09	0.35	17.81	**229.75**	250
Total	252.01	246.78	246.05	255.16	1000

55% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**246.05**	2.54	0.11	1.3	250
Grass	0.91	**225.65**	23.24	0.2	250
Bare Soil	0	15.99	**211.41**	22.6	250
Water	1.65	0.2	18.7	**229.45**	250
Total	248.61	244.38	253.46	253.55	1000

60% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**244.34**	3.54	0.14	1.98	250
Grass	0.76	**223.2**	25.75	0.29	250
Bare Soil	0	14.18	**215.05**	20.77	250
Water	1.27	0.26	20.44	**228.03**	250
Total	246.37	241.18	261.38	251.07	1000

65% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**243.4**	4.09	0.14	2.37	250
Grass	0.71	**221.8**	27.29	0.2	250
Bare Soil	0	12.82	**218**	19.18	250
Water	1.21	0.23	23.04	**225.52**	250
Total	245.32	238.94	268.47	247.27	1000

70% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**241.2**	4.9	0.13	3.77	250
Grass	0.55	**219.21**	29.9	0.34	250
Bare Soil	0	12.09	**219.8**	18.11	250
Water	1.05	0.23	25.37	**223.35**	250
Total	242.8	236.43	275.2	245.57	1000

75% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**239.07**	3.54	0.16	7.23	250
Grass	0.47	**216**	32.72	0.81	250
Bare Soil	0	11.26	**221.47**	17.27	250
Water	0.89	0.28	27.7	**221.13**	250
Total	240.43	231.08	282.05	246.44	1000

80% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**232.71**	1.81	0.14	15.34	250
Grass	0.31	**212.59**	35.75	1.35	250
Bare Soil	0	10.18	**223.73**	16.09	250
Water	0.84	0.17	30.84	**218.15**	250
Total	233.86	224.75	290.46	250.93	1000

85% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**224.01**	1.23	0.17	24.59	250
Grass	0.16	**208.28**	39.65	1.91	250
Bare Soil	0	9.13	**226.07**	14.8	250
Water	0.5	0.21	35.11	**214.18**	250
Total	224.67	218.85	301	255.48	1000

90% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**213.95**	0.63	0.17	35.25	250
Grass	0.13	**206.22**	41.57	2.08	250
Bare Soil	0	8.44	**227.56**	14	250
Water	0.33	0.26	37.99	**211.42**	250
Total	214.41	215.55	307.29	262.75	1000

95% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**203.55**	0.16	0.17	46.12	250
Grass	0.09	**199.96**	46.84	3.11	250
Bare Soil	0	6.74	**230.95**	12.31	250
Water	0.17	0.02	44.78	**205.03**	250
Total	203.81	206.88	322.74	266.57	1000

100% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**189.85**	0.08	0.21	59.86	250
Grass	0.08	**189.89**	57.06	2.97	250
Bare Soil	0	3.44	**236.29**	10.27	250
Water	0.07	0	51.96	**197.97**	250
Total	190	193.41	345.52	271.07	1000

SOURCE: Author.

Table B.13 - Pixel-based Baseline Classification - Confusion Matrix for SVM-OAO Setup 4.

Pixel-based Baseline Classification. Monte Carlo Simulation using SVM-OAO
Setup 4 Train: [prop, 100%]
Test: [100%]

50% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**241.32**	3.73	1	3.95	250
Grass	0.77	**249**	0.22	0.01	250
Bare Soil	2.91	111.24	**34.88**	100.97	250
Water	0.21	0	0.12	**249.67**	250
Total	245.21	363.97	36.22	354.6	1000

55% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**241.35**	3.43	1.74	3.48	250
Grass	0.32	**248.83**	0.84	0.01	250
Bare Soil	2.69	89.42	**73.6**	84.29	250
Water	0.53	0	0	**249.47**	250
Total	244.89	341.68	76.18	337.25	1000

60% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**240.73**	2.91	2.99	3.37	250
Grass	0.15	**246.37**	3.48	0	250
Bare Soil	2.28	55.54	**133.13**	59.05	250
Water	0.06	0	0.51	**249.43**	250
Total	243.22	304.82	140.11	311.85	1000

65% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**240.78**	2.86	3.63	2.73	250
Grass	0.1	**244.55**	5.35	0	250
Bare Soil	1.9	45.35	**176.44**	26.31	250
Water	0.17	0	6.72	**243.11**	250
Total	242.95	292.76	192.14	272.15	1000

70% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**240.24**	2.81	4.32	2.63	250
Grass	0.02	**243.88**	6.1	0	250
Bare Soil	1.56	42.17	**185.04**	21.23	250
Water	0.15	0	10.88	**238.97**	250
Total	241.97	288.86	206.34	262.83	1000

75% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**239.65**	2.64	5.22	2.49	250
Grass	0.01	**241.12**	8.87	0	250
Bare Soil	1.47	34.16	**194.94**	19.43	250
Water	0.11	0	13.09	**236.8**	250
Total	241.24	277.92	222.12	258.72	1000

80% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**238.52**	2.55	6.48	2.45	250
Grass	0.03	**240.09**	9.88	0	250
Bare Soil	1.12	32.17	**198.62**	18.09	250
Water	0.17	0	15.55	**234.28**	250
Total	239.84	274.81	230.53	254.82	1000

85% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**237.08**	2.54	8	2.38	250
Grass	0.39	**240.13**	9.48	0	250
Bare Soil	0.87	28.71	**202.34**	18.08	250
Water	0.13	0	15.27	**234.6**	250
Total	238.47	271.38	235.09	255.06	1000

90% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**234.83**	2.53	10.11	2.53	250
Grass	0.01	**238.42**	11.57	0	250
Bare Soil	0.84	26.78	**205.75**	16.63	250
Water	0.14	0	17.65	**232.21**	250
Total	235.82	267.73	245.08	251.37	1000

95% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**236.88**	2.44	8.04	2.64	250
Grass	0.03	**236.58**	13.39	0	250
Bare Soil	0.84	22.78	**210.15**	16.23	250
Water	0.05	0	18.82	**231.13**	250
Total	237.8	261.8	250.4	250	1000

100% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**230.58**	2.83	14.06	2.53	250
Grass	0	**240.3**	9.7	0	250
Bare Soil	0.68	28.92	**205.29**	15.11	250
Water	0	0	21.97	**228.03**	250
Total	231.26	272.05	251.02	245.67	1000

SOURCE: Author.

Table B.14 - Pixel-based Baseline Classification - Confusion Matrix for KNN-5 Setup 4.

Pixel-based Baseline Classification. Monte Carlo Simulation using KNN-5
Setup 4 Train: [prop, 100%]
Test: [100%]

50% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0	0.52	**248.69**	0.79	250
Water	1.62	0.02	0.12	**248.24**	250
Total	251.62	250.54	248.81	249.03	1000

55% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0	0.27	**249.17**	0.56	250
Water	0.98	0.05	0.07	**248.9**	250
Total	250.98	250.32	249.24	249.46	1000

60% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0	0.19	**249.53**	0.28	250
Water	0.66	0.02	0.04	**249.28**	250
Total	250.66	250.21	249.57	249.56	1000

65% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0	0.21	**249.62**	0.17	250
Water	0.51	0.01	0.11	**249.37**	250
Total	250.51	250.22	249.73	249.54	1000

70% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**249.99**	0.01	0	250
Bare Soil	0	0.18	**249.66**	0.16	250
Water	0.54	0.05	0.07	**249.34**	250
Total	250.54	250.22	249.74	249.5	1000

75% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0	0.13	**249.72**	0.15	250
Water	0.32	0.02	0.07	**249.59**	250
Total	250.32	250.15	249.79	249.74	1000

80% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0	0.06	**249.82**	0.12	250
Water	0.13	0	0.06	**249.81**	250
Total	250.13	250.06	249.88	249.93	1000

85% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0	0.02	**249.91**	0.07	250
Water	0.04	0.02	0.07	**249.87**	250
Total	250.04	250.04	249.98	249.94	1000

90% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0	0	**249.89**	0.11	250
Water	0.02	0	0.06	**249.92**	250
Total	250.02	250	249.95	250.03	1000

95% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0	0.03	**249.96**	0.01	250
Water	0	0	0.07	**249.93**	250
Total	250	250.03	250.03	249.94	1000

100% — MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**249.95**	0.05	0	250
Bare Soil	0	0	**249.72**	0.28	250
Water	0	0	0.07	**249.93**	250
Total	250	249.95	249.84	250.21	1000

SOURCE: Author.

Table B.15 - Pixel-based Baseline Classification - Confusion Matrix for SVM-OAO Setup 5.

Pixel-based Baseline Classification. Monte Carlo Simulation using SVM-OAO

Setup 5 Train: [prop, prop + 5%[

Test: [prop, 100%]

50-55%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**241.35**	3.7	0.98	3.97	250
Grass	0.82	**248.97**	0.2	0.01	250
Bare Soil	2.9	111.2	**34.29**	101.61	250
Water	0.27	0	0.04	**249.69**	250
Total	245.34	363.87	35.51	355.28	1000

55-60%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**241.49**	3.03	1.82	3.66	250
Grass	0.12	**249.61**	0.27	0	250
Bare Soil	2.02	87.23	**78.29**	82.46	250
Water	0.53	0	0	**249.47**	250
Total	244.16	339.87	80.38	335.59	1000

60-65%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**240.85**	2.63	2.98	3.54	250
Grass	0.03	**249.24**	0.73	0	250
Bare Soil	1.61	42.19	**153.87**	52.33	250
Water	0.04	0	0.05	**249.91**	250
Total	242.53	294.06	157.63	305.78	1000

65-70%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**241.38**	2.36	3.37	2.89	250
Grass	0	**249.52**	0.48	0	250
Bare Soil	1.11	24.81	**209.91**	14.17	250
Water	0.2	0	0.98	**248.82**	250
Total	242.69	276.69	214.74	265.88	1000

70-75%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**240.98**	2.75	3.46	2.81	250
Grass	0	**249.65**	0.35	0	250
Bare Soil	0.77	17.33	**225.14**	6.76	250
Water	0.26	0	1.39	**248.35**	250
Total	242.01	269.73	230.34	257.92	1000

75-80%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**241.08**	2.53	3.77	2.62	250
Grass	0	**249.67**	0.33	0	250
Bare Soil	0.81	8.43	**236.89**	3.87	250
Water	0.02	0	1	**248.98**	250
Total	241.91	260.63	241.99	255.47	1000

80-85%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**241.27**	2.37	3.87	2.49	250
Grass	0	**249.72**	0.28	0	250
Bare Soil	0.48	4.8	**242.22**	2.5	250
Water	0.2	0	0.61	**249.19**	250
Total	241.95	256.89	246.98	254.18	1000

85-90%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**241.28**	2.57	3.5	2.65	250
Grass	0	**249.84**	0.16	0	250
Bare Soil	0.26	1.69	**246.62**	1.43	250
Water	0.24	0	0.58	**249.18**	250
Total	241.78	254.1	250.86	253.26	1000

90-95%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**242.54**	1.64	3.41	2.41	250
Grass	0	**249.99**	0.01	0	250
Bare Soil	0.13	0.85	**248.4**	0.62	250
Water	0.33	0	0.69	**248.98**	250
Total	243	252.48	252.51	252.01	1000

95-100%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**244.64**	1.37	2.54	1.45	250
Grass	0	**250**	0	0	250
Bare Soil	0.14	0.2	**249.52**	0.14	250
Water	0.26	0	0.53	**249.21**	250
Total	245.04	251.57	252.59	250.8	1000

100%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**250**	0	0	0	250
Grass	0	**250**	0	0	250
Bare Soil	0.14	0	**249.86**	0	250
Water	0	0	0.43	**249.57**	250
Total	250.14	250	250.29	249.57	1000

SOURCE: Author.

Table B.16 - Pixel-based Baseline Classification - Confusion Matrix for KNN-5 Setup 5.

Pixel-based Baseline Classification. Monte Carlo Simulation using KNN-5
Setup 5 Train: [prop, prop + 5%[
Test: [prop, 100%]

50-55%

MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**241.23**	2.98	2.54	3.25	250
Grass	2.1	**226.91**	20.98	0.01	250
Bare Soil	2.99	68.48	**112.37**	66.16	250
Water	1.18	0	28.62	**220.2**	250
Total	247.5	298.37	164.51	289.62	1000

55-60%

MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**241.37**	2.37	3.08	3.18	250
Grass	0.08	**241.19**	8.73	0	250
Bare Soil	1.74	24.41	**192.89**	30.96	250
Water	1.68	0	5.52	**242.8**	250
Total	244.87	267.97	210.22	276.94	1000

60-65%

MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**241**	2.29	3.65	3.06	250
Grass	0.01	**245.72**	4.27	0	250
Bare Soil	1.98	8.85	**224.82**	14.35	250
Water	0.07	0	3.54	**246.39**	250
Total	243.06	256.86	236.28	263.8	1000

65-70%

MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**241.37**	2.02	3.89	2.72	250
Grass	0.04	**248.11**	1.85	0	250
Bare Soil	0.96	3.74	**239.84**	5.46	250
Water	0.56	0	1.72	**247.72**	250
Total	242.93	253.87	247.3	255.9	1000

70-75%

MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**240.82**	2.22	4.23	2.73	250
Grass	0	**248.62**	1.38	0	250
Bare Soil	0.58	2.42	**244.37**	2.63	250
Water	0.01	0	1.13	**248.86**	250
Total	241.41	253.26	251.11	254.22	1000

75-80%

MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**241.04**	2.12	4.35	2.49	250
Grass	0.01	**249.24**	0.75	0	250
Bare Soil	0.47	1.14	**247.16**	1.23	250
Water	0.02	0	0.91	**249.07**	250
Total	241.54	252.5	253.17	252.79	1000

80-85%

MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**241.32**	1.96	4.14	2.58	250
Grass	0.14	**249.07**	0.79	0	250
Bare Soil	0.58	0.58	**248.39**	0.45	250
Water	0.15	0	0.78	**249.07**	250
Total	242.19	251.61	254.1	252.1	1000

85-90%

MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**241.7**	1.86	3.91	2.53	250
Grass	0	**249.37**	0.63	0	250
Bare Soil	0.51	0.37	**248.48**	0.64	250
Water	0.01	0	0.47	**249.52**	250
Total	242.22	251.6	253.49	252.69	1000

90-95%

MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**242.54**	1.43	3.73	2.3	250
Grass	0	**249.51**	0.49	0	250
Bare Soil	0.34	0.04	**249.32**	0.3	250
Water	0.01	0	0.48	**249.51**	250
Total	242.89	250.98	254.02	252.11	1000

95-100%

MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**244.63**	1.28	2.56	1.53	250
Grass	0.03	**249.9**	0.07	0	250
Bare Soil	0.21	0	**249.72**	0.07	250
Water	0	0	0.33	**249.67**	250
Total	244.87	251.18	252.68	251.27	1000

100%

MEAN CONFUSION MATRIX

CLASS	Water	Forest	Crops	Bare Soil	Total
Forest	**250**	0	0	0	250
Grass	0	**249.56**	0.44	0	250
Bare Soil	0.14	0	**249.86**	0	250
Water	0	0	0.26	**249.74**	250
Total	250.14	249.56	250.56	249.74	1000

SOURCE: Author.

Table B.17 - Pixel-based Baseline Classification - Confusion Matrix for SVM-OAO Setup 6.

Pixel-based Baseline Classification. Monte Carlo Simulation using SVM-OAO

Setup 6 Train: [prop, 100%]

Test: [prop, 100%]

50%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**241.32**	3.73	1	3.95	250
Grass	0.77	**249**	0.22	0.01	250
Bare Soil	2.91	111.24	**34.88**	100.97	250
Water	0.21	0	0.12	**249.67**	250
Total	245.21	363.97	36.22	354.6	1000

55%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**241.35**	3.43	1.74	3.48	250
Grass	0.32	**248.83**	0.84	0.01	250
Bare Soil	2.69	89.42	**73.6**	84.29	250
Water	0.53	0	0	**249.47**	250
Total	244.89	341.68	76.18	337.25	1000

60%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**240.73**	2.91	2.99	3.37	250
Grass	0.15	**246.37**	3.48	0	250
Bare Soil	2.28	55.54	**133.13**	59.05	250
Water	0.06	0	0.51	**249.43**	250
Total	243.22	304.82	140.11	311.85	1000

65%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**240.78**	2.86	3.63	2.73	250
Grass	0.1	**244.55**	5.35	0	250
Bare Soil	1.9	45.35	**176.44**	26.31	250
Water	0.17	0	6.72	**243.11**	250
Total	242.95	292.76	192.14	272.15	1000

70%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**240.24**	2.81	4.32	2.63	250
Grass	0.02	**243.88**	6.1	0	250
Bare Soil	1.56	42.17	**185.04**	21.23	250
Water	0.15	0	10.88	**238.97**	250
Total	241.97	288.86	206.34	262.83	1000

75%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**239.65**	2.64	5.22	2.49	250
Grass	0.01	**241.12**	8.87	0	250
Bare Soil	1.47	34.16	**194.94**	19.43	250
Water	0.11	0	13.09	**236.8**	250
Total	241.24	277.92	222.12	258.72	1000

80%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**238.52**	2.55	6.48	2.45	250
Grass	0.03	**240.09**	9.88	0	250
Bare Soil	1.12	32.17	**198.62**	18.09	250
Water	0.17	0	15.55	**234.28**	250
Total	239.84	274.81	230.53	254.82	1000

85%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**237.08**	2.54	8	2.38	250
Grass	0.39	**240.13**	9.48	0	250
Bare Soil	0.87	28.71	**202.34**	18.08	250
Water	0.13	0	15.27	**234.6**	250
Total	238.47	271.38	235.09	255.06	1000

90%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**234.83**	2.53	10.11	2.53	250
Grass	0.01	**238.42**	11.57	0	250
Bare Soil	0.84	26.78	**205.75**	16.63	250
Water	0.14	0	17.65	**232.21**	250
Total	235.82	267.73	245.08	251.37	1000

95%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**236.88**	2.44	8.04	2.64	250
Grass	0.03	**236.58**	13.39	0	250
Bare Soil	0.84	22.78	**210.15**	16.23	250
Water	0.05	0	18.82	**231.13**	250
Total	237.8	261.8	250.4	250	1000

100%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**230.58**	2.83	14.06	2.53	250
Grass	0	**240.3**	9.7	0	250
Bare Soil	0.68	28.92	**205.29**	15.11	250
Water	0	0	21.97	**228.03**	250
Total	231.26	272.05	251.02	245.67	1000

SOURCE: Author.

Table B.18 - Pixel-based Baseline Classification - Confusion Matrix for KNN-5 Setup 6.

Pixel-based Baseline Classification. Monte Carlo Simulation using KNN-5

Setup 6 Train: [prop, 100%]

Test: [prop, 100%]

50%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**240.96**	2.47	3.92	2.65	250
Grass	0.58	**232.32**	17.1	0	250
Bare Soil	2.23	20.61	**207.23**	19.93	250
Water	0.11	0	12.3	**237.59**	250
Total	243.88	255.4	240.55	260.17	1000

55%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**241.1**	2.11	3.96	2.83	250
Grass	0.25	**239.7**	10.05	0	250
Bare Soil	1.44	13.76	**220.45**	14.35	250
Water	0.07	0	7.33	**242.6**	250
Total	242.86	255.57	241.79	259.78	1000

60%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**241.01**	2.11	4.05	2.83	250
Grass	0.25	**243.97**	5.78	0	250
Bare Soil	1.14	7.5	**232.75**	8.61	250
Water	0.06	0	4.49	**245.45**	250
Total	242.46	253.58	247.07	256.89	1000

65%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**241.38**	1.91	4.01	2.7	250
Grass	0.07	**246.88**	3.05	0	250
Bare Soil	0.97	4.43	**239.57**	5.03	250
Water	0	0	3.08	**246.92**	250
Total	242.42	253.22	249.71	254.65	1000

70%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**240.92**	2.28	4.1	2.7	250
Grass	0.07	**247.74**	2.19	0	250
Bare Soil	0.67	3.46	**243.14**	2.73	250
Water	0.02	0	2.07	**247.91**	250
Total	241.68	253.48	251.5	253.34	1000

75%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**241.07**	2.04	4.34	2.55	250
Grass	0.05	**248.92**	1.03	0	250
Bare Soil	0.78	1.66	**245.95**	1.61	250
Water	0.01	0	1.38	**248.61**	250
Total	241.91	252.62	252.7	252.77	1000

80%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**241.36**	2.01	4.17	2.46	250
Grass	0.01	**249.28**	0.71	0	250
Bare Soil	0.61	1.03	**247.53**	0.83	250
Water	0	0	0.76	**249.24**	250
Total	241.98	252.32	253.17	252.53	1000

85%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**241.36**	2.11	3.98	2.55	250
Grass	0	**249.44**	0.56	0	250
Bare Soil	0.34	0.32	**248.75**	0.59	250
Water	0	0	0.66	**249.34**	250
Total	241.7	251.87	253.95	252.48	1000

90%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**242.55**	1.57	3.49	2.39	250
Grass	0.05	**249.6**	0.35	0	250
Bare Soil	0.14	0.09	**249.49**	0.28	250
Water	0	0	0.69	**249.31**	250
Total	242.74	251.26	254.02	251.98	1000

95%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**244.61**	1.29	2.6	1.5	250
Grass	0	**249.74**	0.26	0	250
Bare Soil	0.25	0.07	**249.62**	0.06	250
Water	0	0	0.52	**249.48**	250
Total	244.86	251.1	253	251.04	1000

100%

MEAN CONFUSION MATRIX					
CLASS	**Water**	**Forest**	**Crops**	**Bare Soil**	**Total**
Forest	**250**	0	0	0	250
Grass	0	**249.56**	0.44	0	250
Bare Soil	0.14	0	**249.86**	0	250
Water	0	0	0.26	**249.74**	250
Total	250.14	249.56	250.56	249.74	1000

SOURCE: Author.

B.2 RSS part II- classification results for Region-based baseline classification

In this part, the results Region-based Baseline classification spatial data quality are presented. Firstly the Graphic results are presented (Appendix B.2.1) leading to a visual interpretation of the data. Afterwards the tables with exact data are presented (Appendix B.2.2).

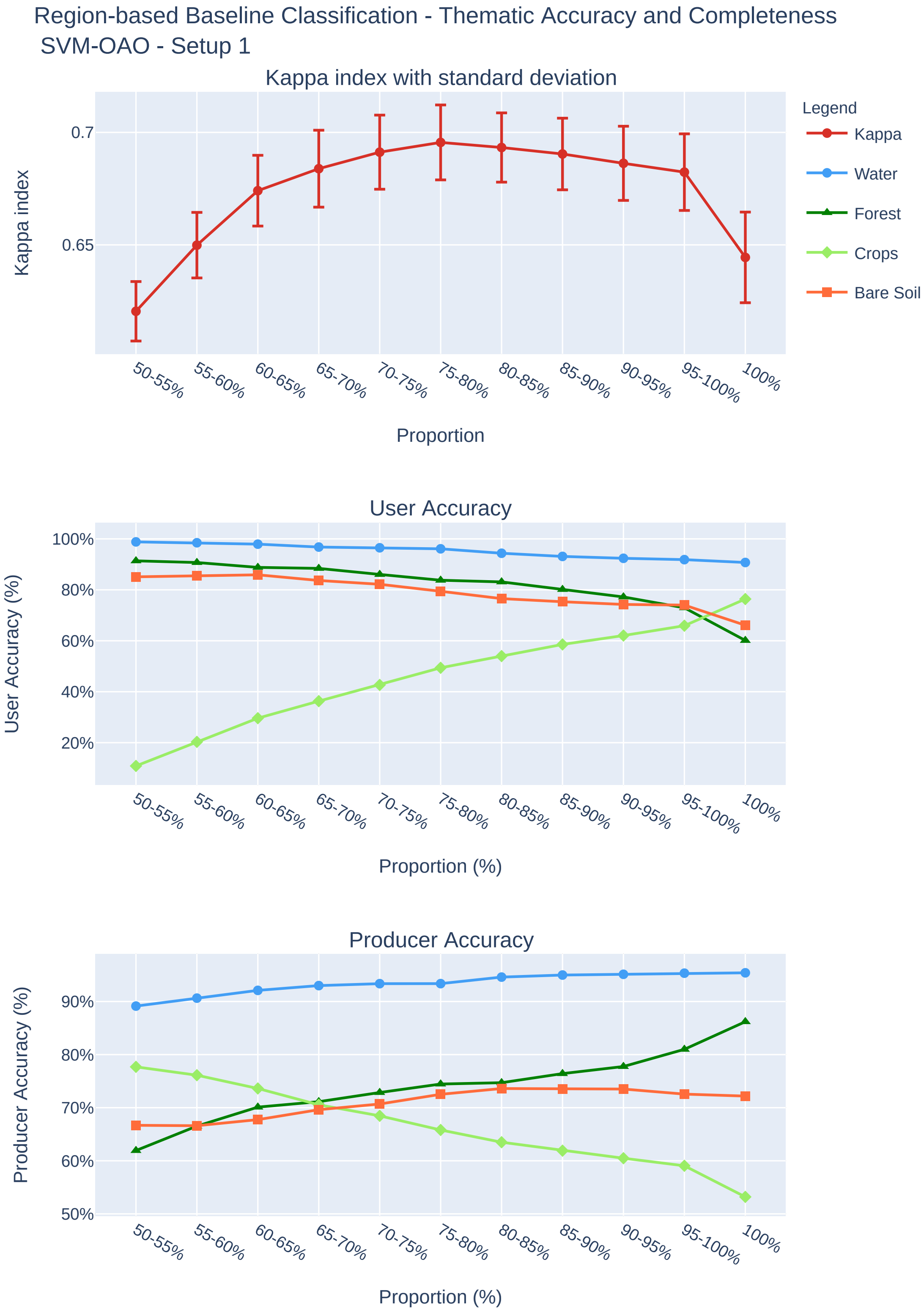
B.2.1 Graphic results

The graphic results present graphs with the mean *kappa* index and the error bars representing their respective standard deviation (std) in the upper area. The middle area presents the mean User Accuracy (UA) whilst the lower area presents the Producer Accuracy (PA).

We note that the vertical scale of all graphs may change according to minimum and maximum value of the presented information.

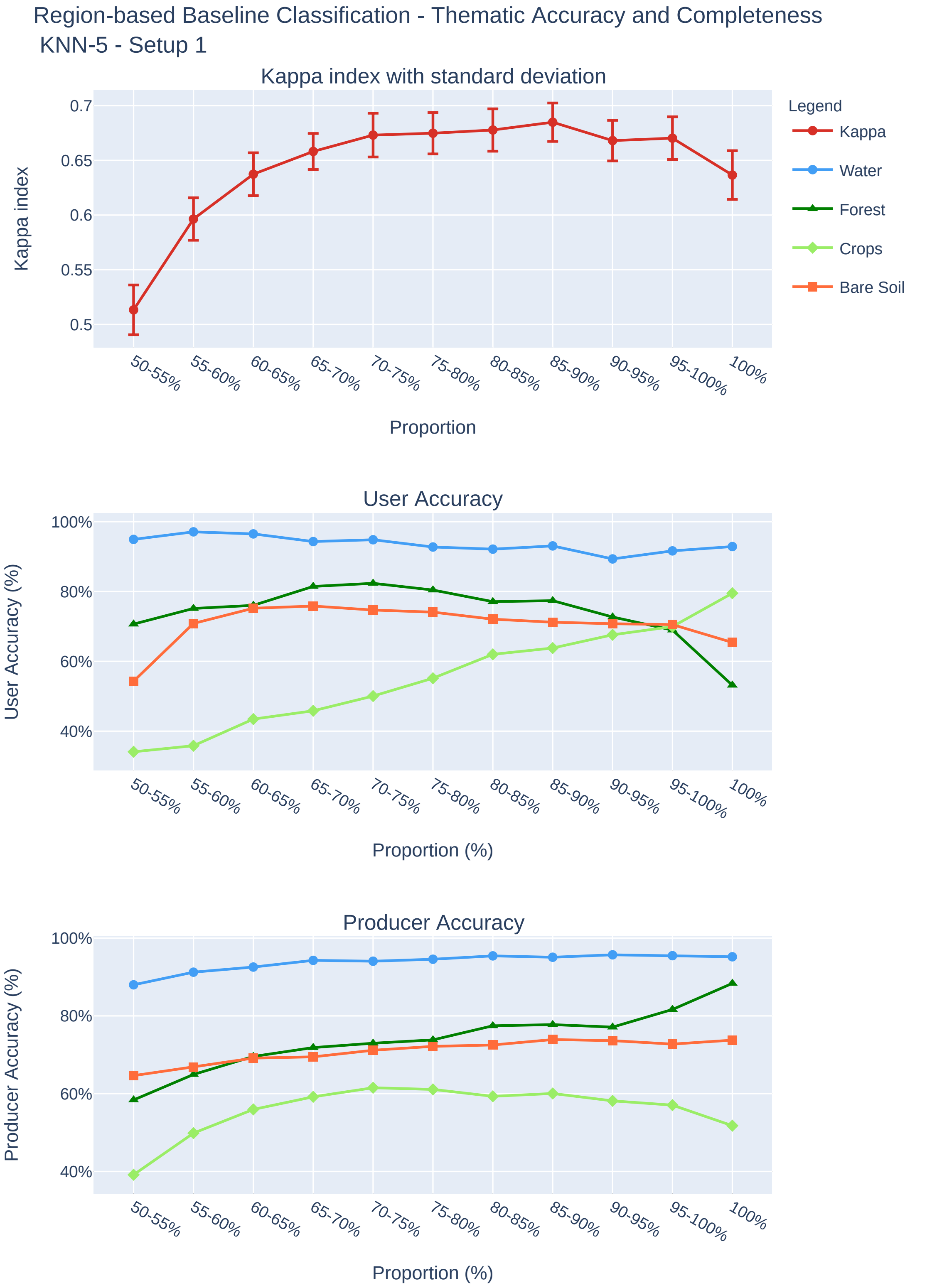
These figures refer to Section 5.2.4, in page 97.

Figure B.13 - Region-based Baseline Classification - Thematic Accuracy and Completeness for SVM-OAO for Setup 1.



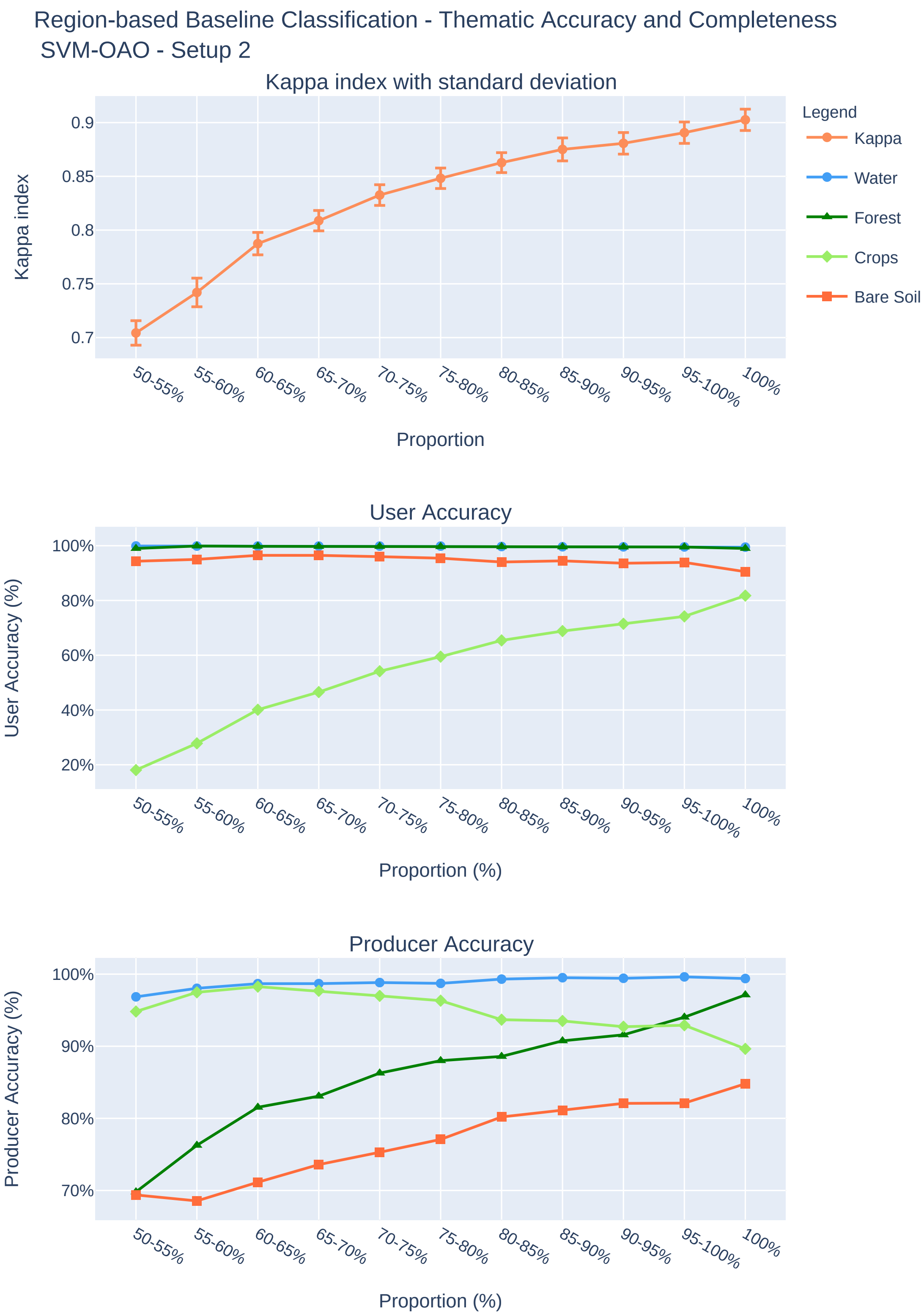
SOURCE: Author.

Figure B.14 - Region-based Baseline Classification - Thematic Accuracy and Completeness for KNN-5 for Setup 1.



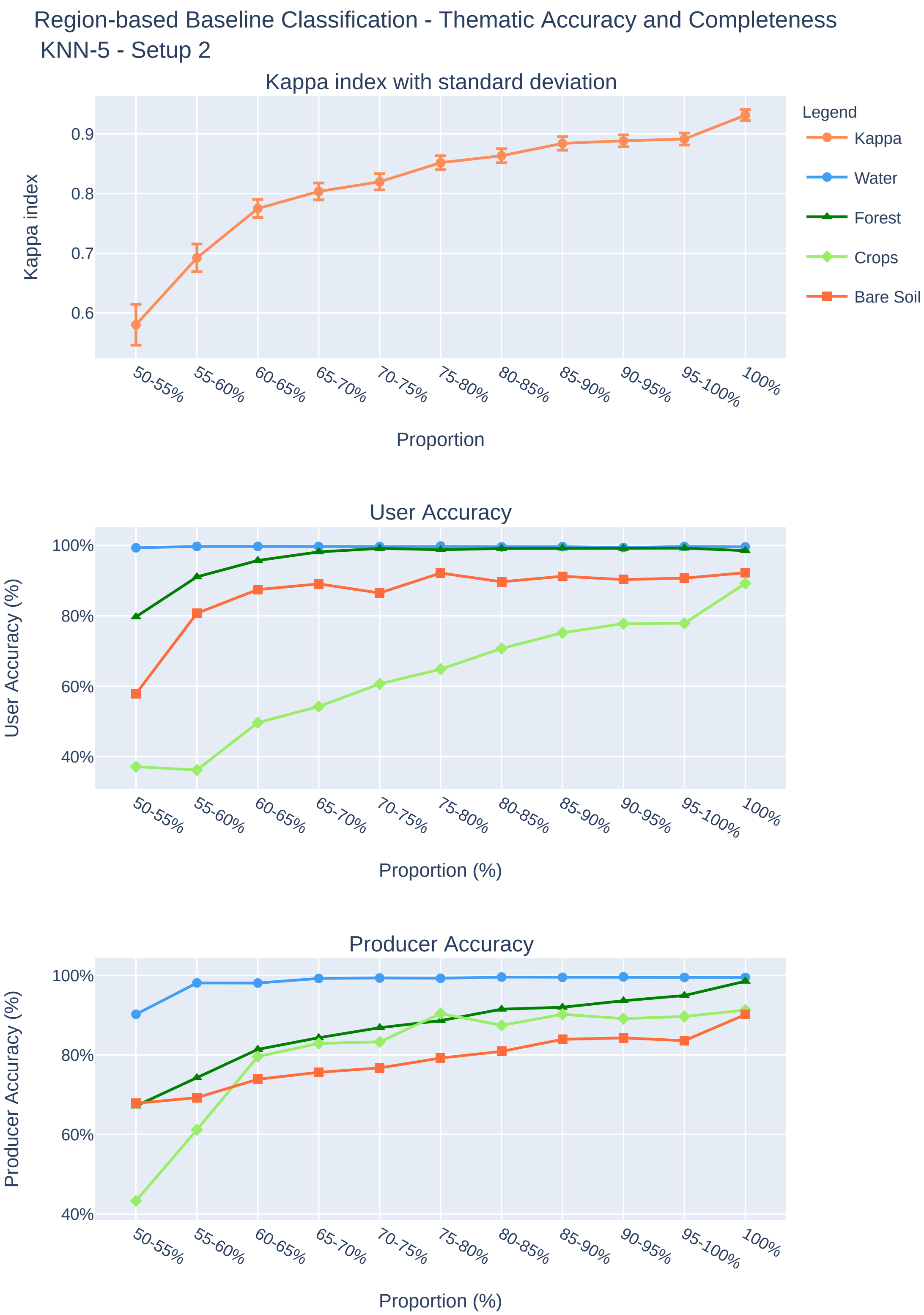
SOURCE: Author.

Figure B.15 - Region-based Baseline Classification - Thematic Accuracy and Completeness for SVM-OAO for Setup 2.



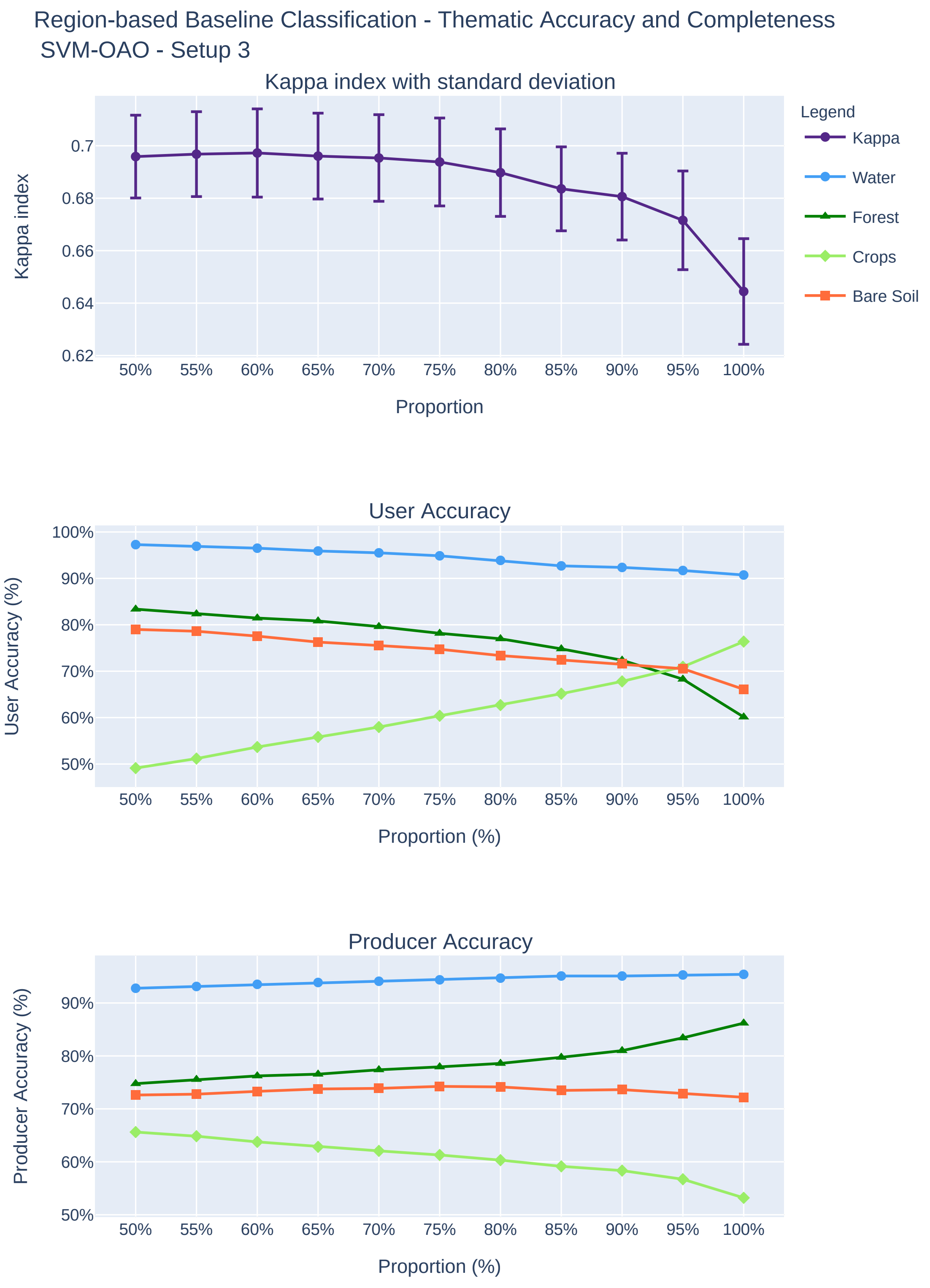
SOURCE: Author.

Figure B.16 - Region-based Baseline Classification - Thematic Accuracy and Completeness for KNN-5 for Setup 2.



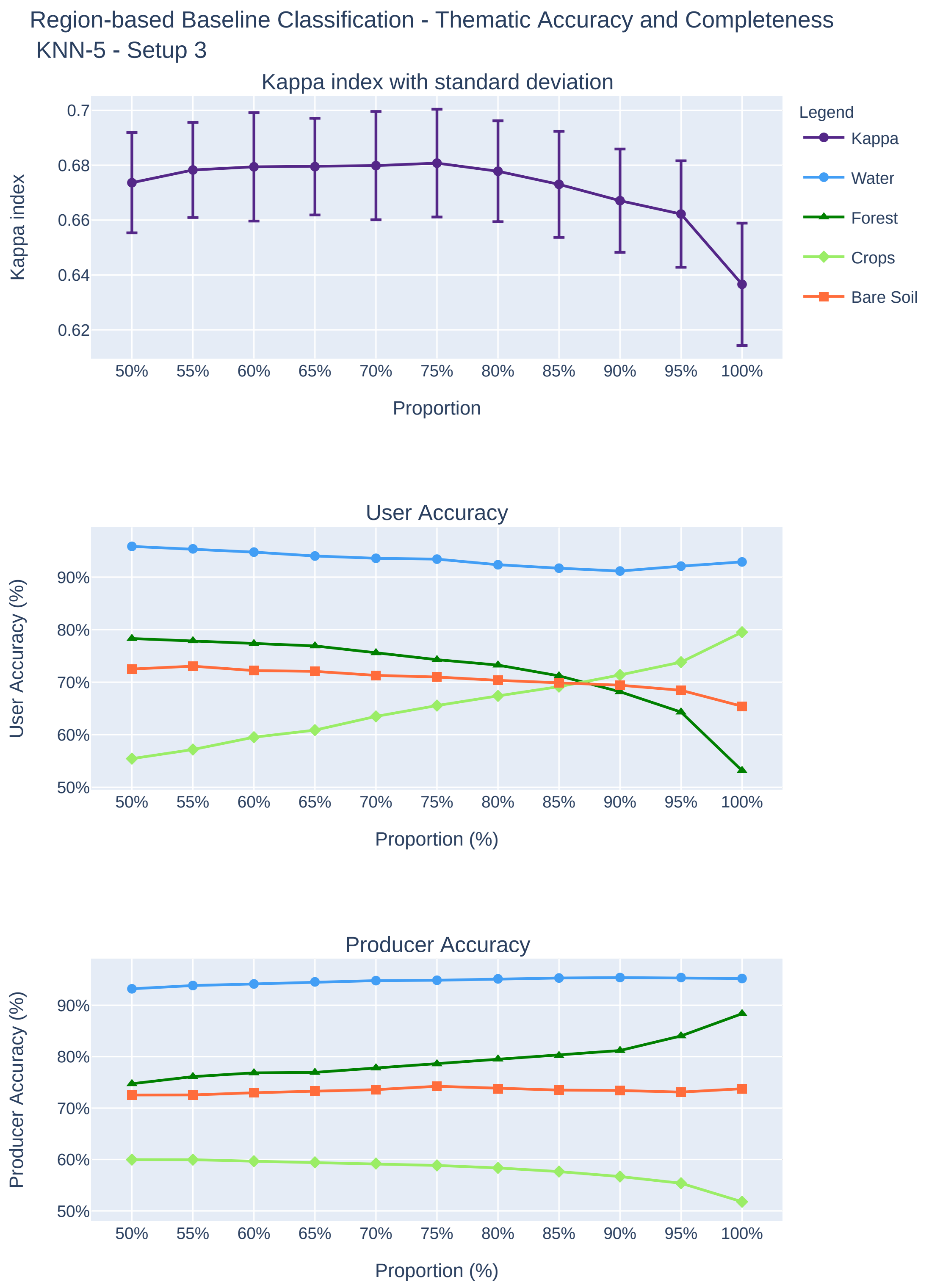
SOURCE: Author.

Figure B.17 - Region-based Baseline Classification - Thematic Accuracy and Completeness for SVM-OAO for Setup 3.



SOURCE: Author.

Figure B.18 - Region-based Baseline Classification - Thematic Accuracy and Completeness for KNN-5 for Setup 3.

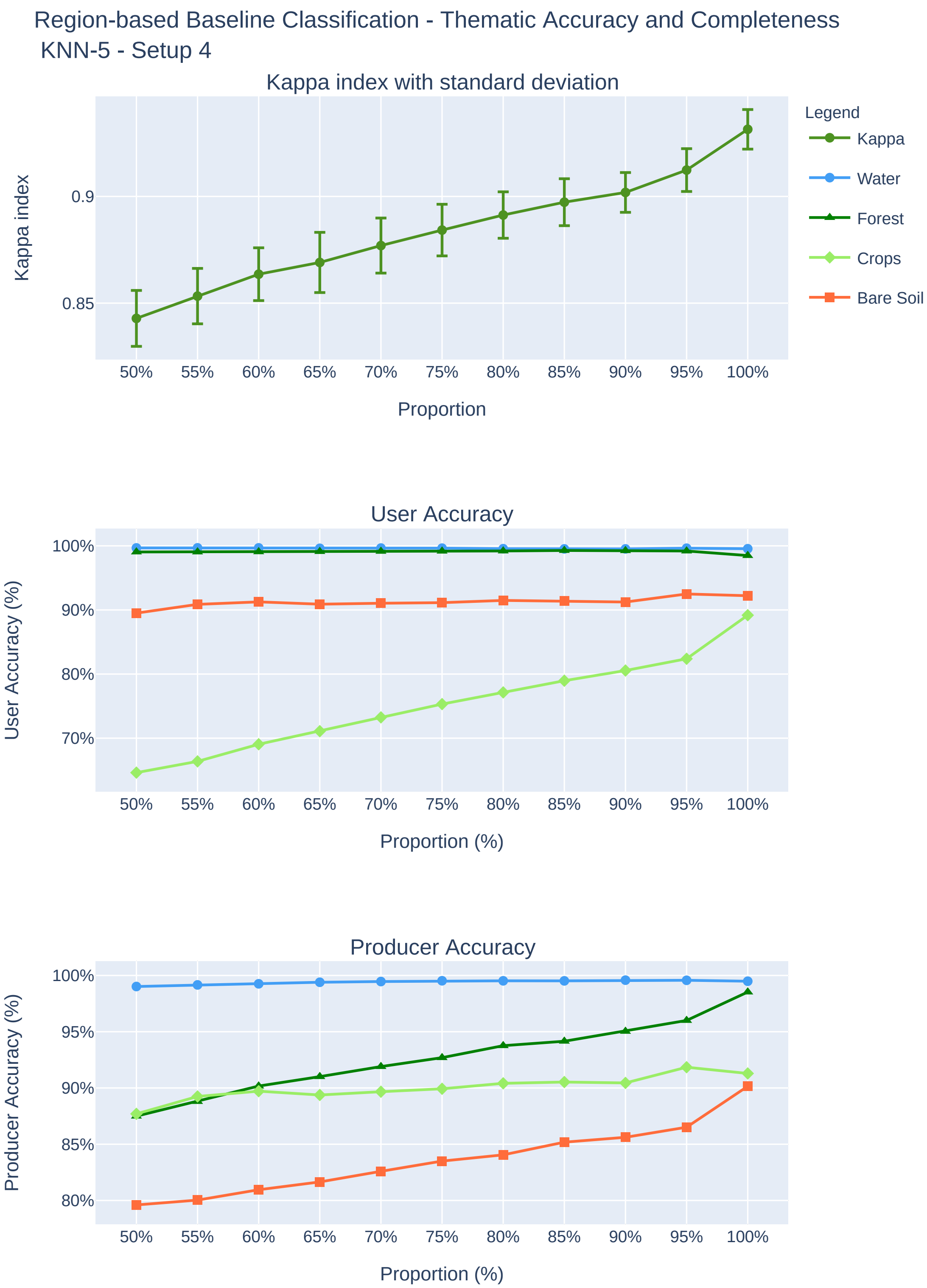


SOURCE: Author.

Figure B.19 - Region-based Baseline Classification - Thematic Accuracy and Completeness for SVM-OAO for Setup 4.

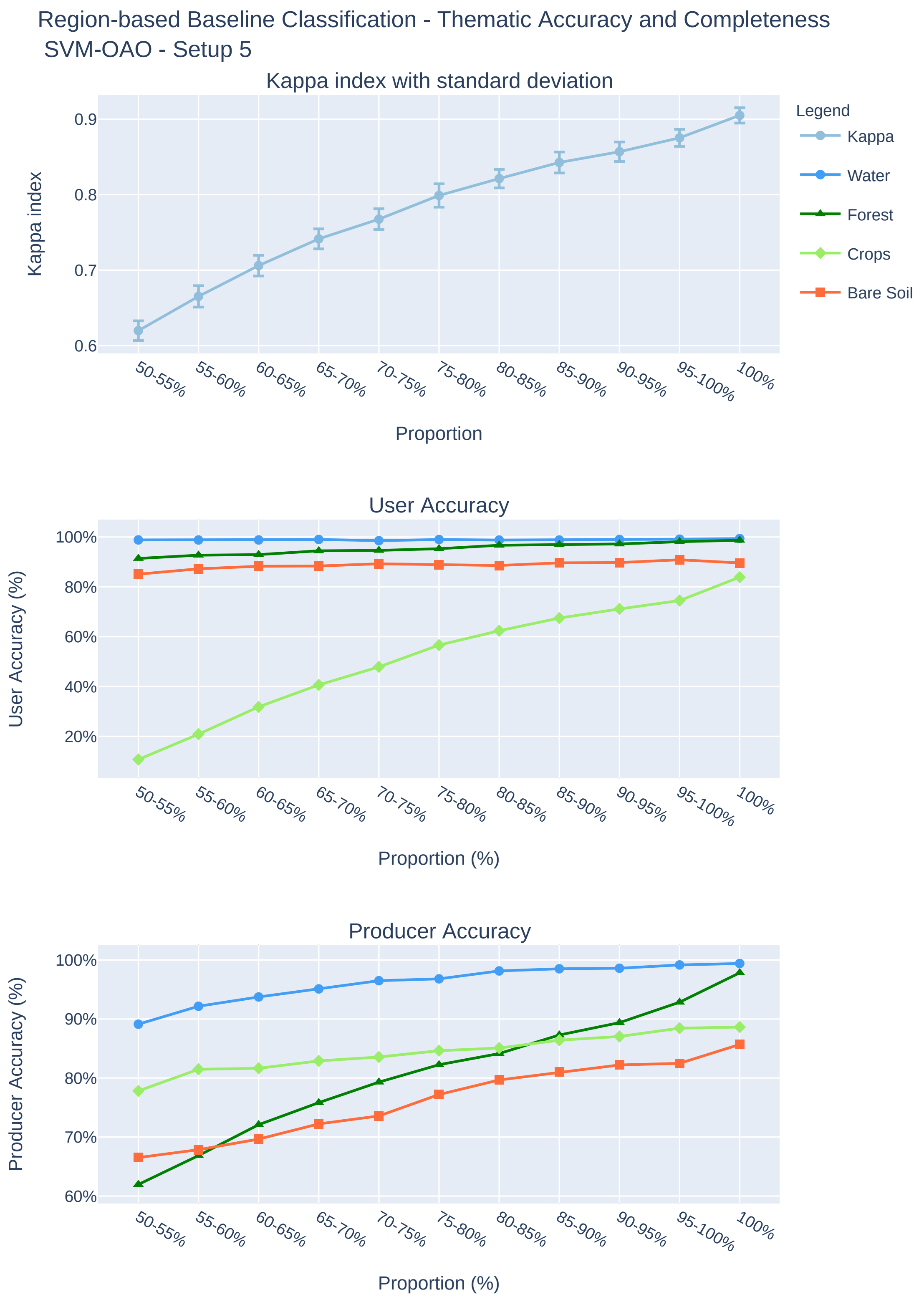
SOURCE: Author.

Figure B.20 - Region-based Baseline Classification - Thematic Accuracy and Completeness for KNN-5 for Setup 4.



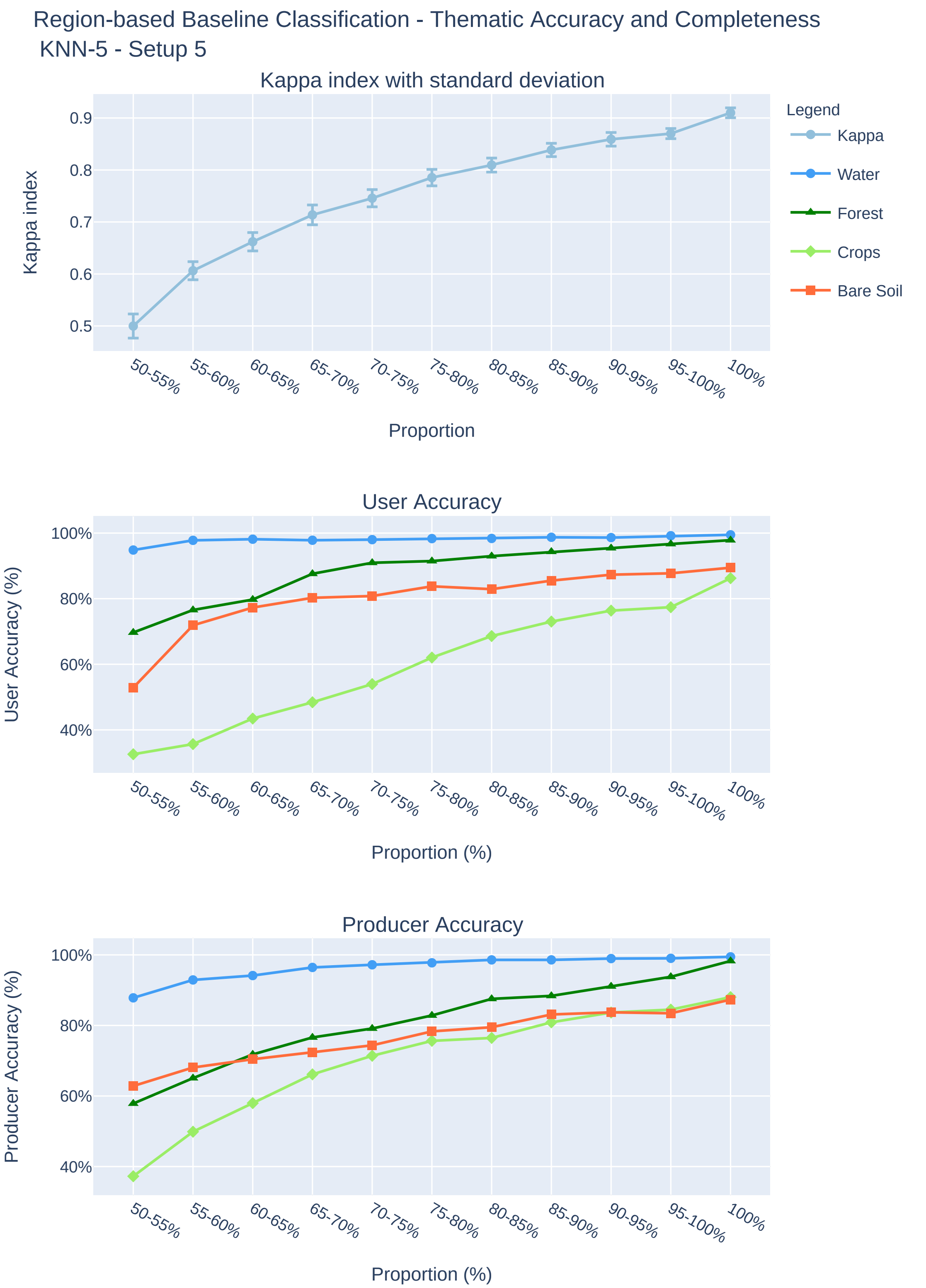
SOURCE: Author.

Figure B.21 - Region-based Baseline Classification - Thematic Accuracy and Completeness for SVM-OAO for Setup 5.



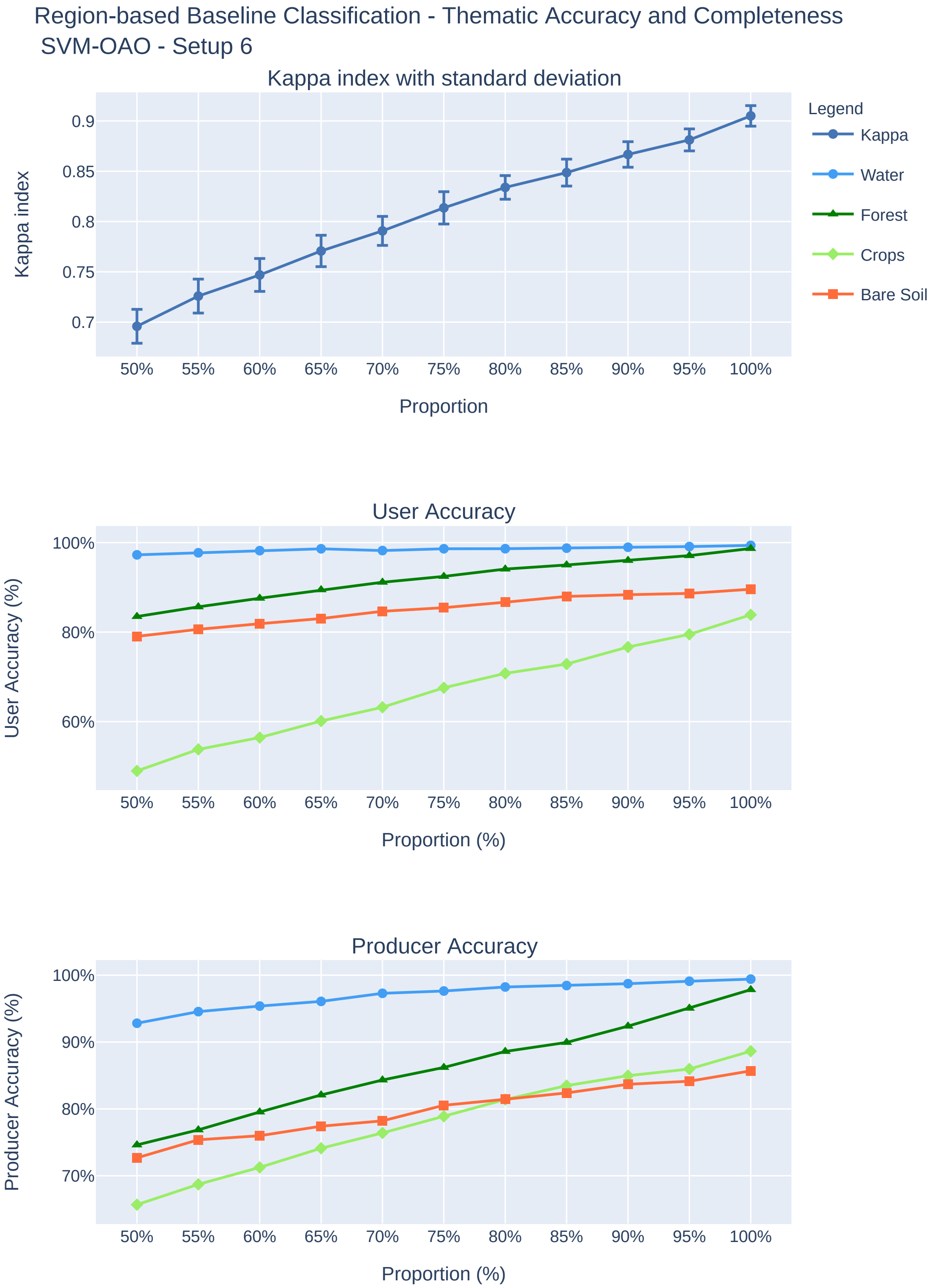
SOURCE: Author.

Figure B.22 - Region-based Baseline Classification - Thematic Accuracy and Completeness for KNN-5 for Setup 5.



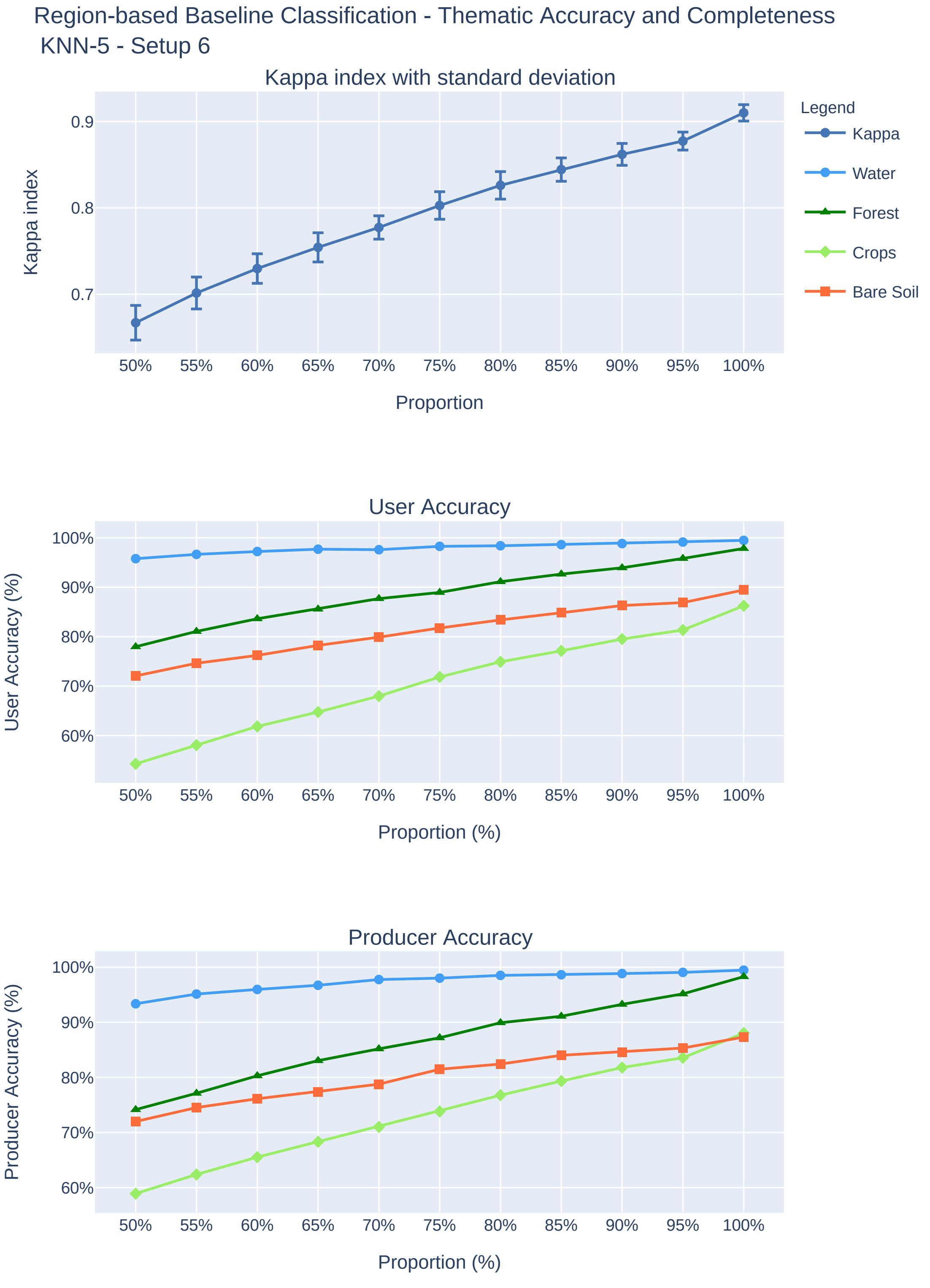
SOURCE: Author.

Figure B.23 - Region-based Baseline Classification - Thematic Accuracy and Completeness for SVM-OAO for Setup 6.



SOURCE: Author.

Figure B.24 - Region-based Baseline Classification - Thematic Accuracy and Completeness for KNN-5 for Setup 6.



SOURCE: Author.

B.2.2 Tabular results

The tabular results for Region-based Baseline Classification are presented in this appendix in two setups per page. Firstly the results regarding SVM-OAO classifier are presented followed by KNN-5 results.

The tables present the mean overall accuracy for the 100 Monte Carlo repetitions with the standard deviation. They also present the User Accuracy (UA) and Producer Accuracy (PA) in percentage for the four used classes.

These tables are in accordance with Section 5.2.4, in page 97.

Table B.19 - Region-based Baseline Classification - Thematic Accuracy and Completeness for SVM-OAO Setups 1 and 2.

Region-based Baseline Classification. Monte Carlo Simulation using SVM-OAO
Setup 1 Train: [prop, prop + 5%[
Test: [50%, 100%]

Interval		**OA**		***Kappa***		**User Accuracy**				**Producer Accuracy**			
Inf	**Sup**	**Mean**	**std**	**Mean**	**std**	**Forest**	**Crops**	**Bare Soil**	**Water**	**Forest**	**Crops**	**Bare Soil**	**Water**
50%	55%	0.7154	0.0099	0.6205	0.0132	91.37%	10.86%	85.10%	98.82%	61.93%	77.69%	66.68%	89.14%
55%	60%	0.7374	0.0109	0.6499	0.0146	90.71%	20.32%	85.46%	98.47%	66.53%	76.12%	66.61%	90.63%
60%	65%	0.7556	0.0118	0.6741	0.0157	88.80%	29.61%	85.89%	97.92%	70.11%	73.62%	67.75%	92.10%
65%	70%	0.7629	0.0128	0.6839	0.0171	88.40%	36.31%	83.66%	96.78%	71.13%	70.54%	69.62%	92.99%
70%	75%	0.7684	0.0124	0.6912	0.0165	86.04%	42.66%	82.17%	96.49%	72.86%	68.46%	70.69%	93.36%
75%	80%	0.7717	0.0125	0.6955	0.0167	83.76%	49.39%	79.42%	96.09%	74.45%	65.80%	72.52%	93.37%
80%	85%	0.7700	0.0115	0.6933	0.0154	83.10%	53.99%	76.54%	94.35%	74.70%	63.49%	73.61%	94.59%
85%	90%	0.7678	0.0119	0.6904	0.0159	80.13%	58.56%	75.32%	93.11%	76.42%	61.90%	73.47%	94.98%
90%	95%	0.7647	0.0124	0.6863	0.0165	77.20%	62.05%	74.26%	92.37%	77.75%	60.48%	73.51%	95.12%
95%	<	0.7618	0.0128	0.6823	0.0170	72.94%	65.91%	74.00%	91.85%	80.99%	59.06%	72.56%	95.33%
100%	100%	0.7333	0.0151	0.6444	0.0201	60.13%	76.38%	66.10%	90.73%	86.19%	53.20%	72.19%	95.40%

Region-based Baseline Classification. Monte Carlo Simulation using SVM-OAO
Setup 2 Train: [prop, prop + 5%[
Test: [100%]

Interval		**OA**		***Kappa***		**User Accuracy**				**Producer Accuracy**			
Inf	**Sup**	**Mean value**	**std**	**Mean value**	**std**	**Forest**	**Crops**	**Bare Soil**	**Water**	**Forest**	**Crops**	**Bare Soil**	**Water**
50%	55%	0.7782	0.0085	0.7043	0.0114	99.00%	18.07%	94.33%	99.89%	69.80%	94.82%	69.38%	96.85%
55%	60%	0.8065	0.0100	0.7420	0.0133	99.89%	27.82%	95.01%	99.88%	76.26%	97.48%	68.55%	98.03%
60%	65%	0.8406	0.0078	0.7874	0.0104	99.78%	40.10%	96.48%	99.86%	81.54%	98.26%	71.14%	98.68%
65%	70%	0.8566	0.0071	0.8087	0.0095	99.75%	46.54%	96.47%	99.86%	83.08%	97.69%	73.59%	98.68%
70%	75%	0.8744	0.0072	0.8326	0.0096	99.75%	54.16%	96.00%	99.86%	86.28%	96.98%	75.30%	98.83%
75%	80%	0.8862	0.0071	0.8482	0.0095	99.67%	59.48%	95.44%	99.86%	88.00%	96.32%	77.08%	98.72%
80%	85%	0.8971	0.0070	0.8628	0.0093	99.66%	65.42%	94.04%	99.71%	88.59%	93.68%	80.20%	99.31%
85%	90%	0.9063	0.0080	0.8751	0.0107	99.58%	68.83%	94.47%	99.63%	90.75%	93.51%	81.07%	99.50%
90%	95%	0.9105	0.0075	0.8807	0.0100	99.53%	71.49%	93.60%	99.60%	91.58%	92.71%	82.08%	99.42%
95%	<	0.9180	0.0075	0.8906	0.0099	99.53%	74.19%	93.90%	99.56%	94.03%	92.91%	82.11%	99.62%
100%	100%	0.9269	0.0074	0.9026	0.0099	99.01%	81.77%	90.51%	99.48%	97.12%	89.64%	84.80%	99.39%

SOURCE: Author.

Table B.20 - Region-based Baseline Classification - Thematic Accuracy and Completeness for KNN-5 Setups 1 and 2.

Region-based Baseline Classification. Monte Carlo Simulation using KNN-5
Setup 1 Train: [prop, prop + 5%[
Test: [50%, 100%]

Interval		OA		*Kappa*		User Accuracy				Producer Accuracy			
Inf	**Sup**	**Mean**	**std**	**Mean**	**std**	**Forest**	**Crops**	**Bare Soil**	**Water**	**Forest**	**Crops**	**Bare Soil**	**Water**
50%	55%	0.6350	0.0171	0.5133	0.0227	70.67%	34.09%	54.31%	94.94%	58.39%	39.16%	64.64%	87.99%
55%	60%	0.6973	0.0146	0.5964	0.0194	75.15%	35.82%	70.85%	97.10%	64.94%	49.86%	66.84%	91.24%
60%	65%	0.7280	0.0147	0.6374	0.0196	76.04%	43.45%	75.21%	96.51%	69.57%	55.95%	69.15%	92.56%
65%	70%	0.7436	0.0123	0.6582	0.0164	81.46%	45.83%	75.84%	94.32%	71.85%	59.20%	69.47%	94.28%
70%	75%	0.7548	0.0150	0.6731	0.0200	82.36%	50.05%	74.70%	94.83%	72.97%	61.52%	71.17%	94.06%
75%	80%	0.7561	0.0142	0.6748	0.0189	80.43%	55.16%	74.10%	92.76%	73.83%	61.09%	72.17%	94.56%
80%	85%	0.7583	0.0145	0.6778	0.0194	77.10%	62.01%	72.08%	92.14%	77.46%	59.33%	72.52%	95.43%
85%	90%	0.7637	0.0132	0.6849	0.0176	77.38%	63.81%	71.21%	93.07%	77.76%	60.05%	73.92%	95.08%
90%	95%	0.7511	0.0139	0.6681	0.0186	72.71%	67.60%	70.77%	89.34%	77.12%	58.15%	73.63%	95.70%
95%	< 100%	0.7527	0.0147	0.6703	0.0195	68.97%	69.95%	70.53%	91.64%	81.65%	57.06%	72.75%	95.48%
100%	100%	0.7275	0.0167	0.6366	0.0223	53.16%	79.52%	65.41%	92.89%	88.35%	51.77%	73.77%	95.20%

Region-based Baseline Classification. Monte Carlo Simulation using KNN-5
Setup 2 Train: [prop, prop + 5%[
Test: [100%]

Interval		OA		*Kappa*		User Accuracy				Producer Accuracy			
Inf	**Sup**	**Mean**	**std**	**Mean**	**std**	**Forest**	**Crops**	**Bare Soil**	**Water**	**Forest**	**Crops**	**Bare Soil**	**Water**
50%	55%	0.6851	0.0257	0.5802	0.0343	79.70%	37.18%	57.86%	99.30%	67.15%	43.27%	67.82%	90.21%
55%	60%	0.7691	0.0174	0.6922	0.0232	91.04%	36.22%	80.71%	99.68%	74.25%	61.16%	69.25%	98.10%
60%	65%	0.8312	0.0114	0.7750	0.0152	95.70%	49.68%	87.44%	99.68%	81.39%	79.58%	73.87%	98.08%
65%	70%	0.8527	0.0106	0.8036	0.0141	98.13%	54.24%	89.04%	99.66%	84.32%	82.88%	75.63%	99.23%
70%	75%	0.8648	0.0102	0.8197	0.0136	99.12%	60.67%	86.47%	99.66%	86.84%	83.29%	76.72%	99.35%
75%	80%	0.8888	0.0088	0.8518	0.0117	98.75%	64.85%	92.13%	99.80%	88.62%	90.40%	79.21%	99.28%
80%	85%	0.8976	0.0088	0.8634	0.0117	99.09%	70.72%	89.65%	99.57%	91.50%	87.43%	80.89%	99.58%
85%	90%	0.9131	0.0085	0.8841	0.0114	99.22%	75.19%	91.21%	99.61%	91.98%	90.21%	83.92%	99.51%
90%	95%	0.9164	0.0074	0.8885	0.0099	99.14%	77.77%	90.27%	99.36%	93.62%	89.12%	84.27%	99.63%
95%	< 100%	0.9186	0.0076	0.8914	0.0101	99.20%	77.87%	90.68%	99.67%	94.94%	89.66%	83.58%	99.50%
100%	100%	0.9486	0.0070	0.9314	0.0093	98.48%	89.18%	92.22%	99.55%	98.53%	91.29%	90.15%	99.49%

SOURCE: Author.

Table B.21 - Region-based Baseline Classification - Thematic Accuracy and Completeness for SVM-OAO Setups 3 and 4.

Region-based Baseline Classification. Monte Carlo Simulation using SVM-OAO
Setup 3 Train: [prop, 100%]
Test: [50%, 100%]

Threshold	OA		*Kappa*		User Accuracy				Producer Accuracy			
	Mean	**std**	**Mean**	**std**	**Forest**	**Crops**	**Bare Soil**	**Water**	**Forest**	**Crops**	**Bare Soil**	**Water**
50%	0.7719	0.0118	0.6959	0.0158	83.37%	49.12%	79.00%	97.28%	74.76%	65.62%	72.61%	92.78%
55%	0.7726	0.0121	0.6968	0.0162	82.34%	51.17%	78.61%	96.92%	75.56%	64.84%	72.78%	93.12%
60%	0.7729	0.0126	0.6973	0.0168	81.45%	53.66%	77.56%	96.51%	76.23%	63.76%	73.31%	93.52%
65%	0.7721	0.0123	0.6961	0.0164	80.82%	55.83%	76.27%	95.90%	76.55%	62.83%	73.76%	93.87%
70%	0.7715	0.0124	0.6953	0.0165	79.61%	57.96%	75.53%	95.50%	77.38%	62.07%	73.87%	94.10%
75%	0.7704	0.0126	0.6938	0.0168	78.16%	60.39%	74.73%	94.86%	77.93%	61.29%	74.25%	94.37%
80%	0.7673	0.0125	0.6898	0.0167	76.99%	62.70%	73.36%	93.88%	78.58%	60.32%	74.15%	94.69%
85%	0.7627	0.0120	0.6836	0.0160	74.81%	65.15%	72.42%	92.70%	79.74%	59.15%	73.48%	95.10%
90%	0.7605	0.0124	0.6806	0.0165	72.36%	67.80%	71.66%	92.36%	81.00%	58.34%	73.63%	95.09%
95%	0.7537	0.0141	0.6716	0.0188	68.28%	70.96%	70.54%	91.70%	83.41%	56.71%	72.88%	95.29%
100%	0.7333	0.0151	0.6444	0.0201	60.13%	76.38%	66.10%	90.73%	86.19%	53.20%	72.19%	95.40%

Region-based Baseline Classification. Monte Carlo Simulation using SVM-OAO
Setup 4 Train: [prop, 100%]
Test: [100%]

Threshold	OA		*Kappa*		User Accuracy				Producer Accuracy			
	Mean	**std**	**Mean**	**std**	**Forest**	**Crops**	**Bare Soil**	**Water**	**Forest**	**Crops**	**Bare Soil**	**Water**
50%	0.8880	0.0086	0.8507	0.0115	99.68%	60.34%	95.33%	99.86%	88.16%	95.93%	77.52%	98.81%
55%	0.8925	0.0079	0.8566	0.0105	99.65%	62.26%	95.22%	99.85%	88.68%	95.61%	78.36%	98.88%
60%	0.8966	0.0083	0.8621	0.0110	99.65%	62.26%	95.22%	99.85%	88.68%	95.61%	78.36%	98.88%
65%	0.8966	0.0083	0.8621	0.0110	99.64%	64.32%	94.87%	99.80%	89.40%	94.91%	79.08%	98.99%
70%	0.9015	0.0078	0.8687	0.0104	99.60%	66.92%	94.30%	99.78%	90.05%	94.03%	80.24%	99.10%
75%	0.9052	0.0078	0.8736	0.0104	99.56%	68.71%	94.10%	99.70%	90.67%	93.40%	80.96%	99.30%
80%	0.9091	0.0072	0.8788	0.0096	99.52%	70.45%	94.02%	99.64%	91.40%	93.06%	81.59%	99.47%
85%	0.9131	0.0078	0.8841	0.0104	99.53%	72.30%	93.80%	99.61%	92.38%	92.84%	82.10%	99.47%
90%	0.9164	0.0072	0.8885	0.0096	99.53%	74.30%	93.14%	99.60%	93.19%	92.29%	82.73%	99.48%
95%	0.9225	0.0075	0.8967	0.0100	99.51%	76.72%	93.24%	99.54%	95.07%	92.40%	83.00%	99.59%
100%	0.9269	0.0074	0.9026	0.0099	99.01%	81.77%	90.51%	99.48%	97.12%	89.64%	84.80%	99.39%

SOURCE: Author.

Table B.22 - Region-based Baseline Classification - Thematic Accuracy and Completeness for KNN-5 Setups 3 and 4.

Region-based Baseline Classification. Monte Carlo Simulation using KNN-5
Setup 3 Train: [prop, 100%]
Test: [50%, 100%]

Threshold	OA		*Kappa*		User Accuracy				Producer Accuracy			
	Mean	std	Mean	std	Forest	Crops	Bare Soil	Water	Forest	Crops	Bare Soil	Water
50%	0.7552	0.0137	0.6736	0.0183	78.30%	55.44%	72.50%	95.84%	74.73%	59.97%	72.54%	93.20%
55%	0.7587	0.0130	0.6783	0.0173	77.88%	57.18%	73.04%	95.38%	76.12%	59.96%	72.55%	93.83%
60%	0.7596	0.0148	0.6794	0.0198	77.34%	59.52%	72.20%	94.76%	76.85%	59.66%	72.98%	94.14%
65%	0.7596	0.0132	0.6795	0.0176	76.90%	60.87%	72.06%	94.02%	76.94%	59.45%	73.36%	94.54%
70%	0.7599	0.0148	0.6798	0.0197	75.60%	63.48%	71.29%	93.58%	77.79%	59.21%	73.58%	94.79%
75%	0.7606	0.0147	0.6807	0.0196	74.28%	65.54%	70.98%	93.42%	78.59%	58.84%	74.24%	94.86%
80%	0.7583	0.0138	0.6778	0.0184	73.25%	67.39%	70.34%	92.35%	79.52%	58.36%	73.78%	95.13%
85%	0.7548	0.0145	0.6730	0.0193	71.20%	69.15%	69.86%	91.70%	80.21%	57.65%	73.50%	95.29%
90%	0.7503	0.0141	0.6671	0.0188	68.18%	71.37%	69.42%	91.16%	81.19%	56.68%	73.42%	95.38%
95%	0.7467	0.0145	0.6622	0.0194	64.31%	73.81%	68.46%	92.08%	84.03%	55.38%	73.10%	95.38%
100%	0.7275	0.0167	0.6366	0.0223	53.16%	79.52%	65.41%	92.89%	88.35%	51.77%	73.77%	95.20%

Region-based Baseline Classification. Monte Carlo Simulation using KNN-5
Setup 4 Train: [prop, 100%]
Test: [100%]

Threshold	OA		*Kappa*		User Accuracy				Producer Accuracy			
	Mean	std	Mean	std	Forest	Crops	Bare Soil	Water	Forest	Crops	Bare Soil	Water
50%	0.8821	0.0098	0.8429	0.0131	99.04%	64.62%	89.51%	99.68%	87.49%	87.70%	79.60%	99.02%
55%	0.8900	0.0097	0.8533	0.0130	99.05%	66.38%	90.87%	99.69%	88.80%	89.24%	80.03%	99.16%
60%	0.8977	0.0093	0.8636	0.0124	99.07%	69.05%	91.26%	99.68%	90.17%	89.72%	80.95%	99.26%
65%	0.9018	0.0106	0.8691	0.0141	99.11%	71.09%	90.90%	99.62%	91.01%	89.37%	81.65%	99.40%
70%	0.9077	0.0097	0.8770	0.0129	99.14%	73.24%	91.05%	99.66%	91.90%	89.67%	82.59%	99.46%
75%	0.9132	0.0091	0.8842	0.0121	99.17%	75.32%	91.14%	99.64%	92.69%	89.93%	83.50%	99.53%
80%	0.9185	0.0082	0.8913	0.0109	99.20%	77.14%	91.49%	99.56%	93.76%	90.41%	84.05%	99.52%
85%	0.9230	0.0082	0.8973	0.0110	99.28%	78.96%	91.42%	99.52%	94.15%	90.53%	85.18%	99.52%
90%	0.9264	0.0070	0.9019	0.0093	99.25%	80.55%	91.24%	99.52%	95.03%	90.45%	85.62%	99.58%
95%	0.9343	0.9343	0.9123	0.0100	99.20%	82.38%	92.49%	99.63%	96.00%	91.84%	86.51%	99.57%
100%	0.9486	0.0070	0.9314	0.0093	98.48%	89.18%	92.22%	99.55%	98.53%	91.29%	90.15%	99.49%

SOURCE: Author.

Table B.23 - Region-based Baseline Classification - Thematic Accuracy and Completeness for SVM-OAO Setups 5 and 6.

Region-based Baseline Classification. Monte Carlo Simulation using SVM-OAO
Setup 5 Train: [prop, prop + 5%[
Test: [prop, 100%]

Interval		OA		*Kappa*		User Accuracy				Producer Accuracy			
Inf	**Sup**	**Mean**	**std**	**Mean**	**std**	**Forest**	**Crops**	**Bare Soil**	**Water**	**Forest**	**Crops**	**Bare Soil**	**Water**
50%	55%	0.7149	0.0098	0.6199	0.0130	91.38%	10.69%	85.08%	98.82%	61.96%	77.81%	66.52%	89.12%
55%	60%	0.7489	0.0107	0.6652	0.0142	92.70%	20.86%	87.22%	98.80%	66.85%	81.48%	67.85%	92.16%
60%	65%	0.7795	0.0103	0.7060	0.0137	92.93%	31.82%	88.25%	98.81%	72.09%	81.65%	69.65%	93.74%
65%	70%	0.8061	0.0099	0.7415	0.0132	94.45%	40.63%	88.36%	99.00%	75.83%	82.90%	72.23%	95.10%
70%	75%	0.8257	0.0103	0.7675	0.0138	94.62%	47.86%	89.24%	98.53%	79.31%	83.56%	73.57%	96.49%
75%	80%	0.8492	0.0116	0.7989	0.0154	95.29%	56.62%	88.80%	98.98%	82.26%	84.63%	77.20%	96.81%
80%	85%	0.8660	0.0092	0.8213	0.0123	96.68%	62.37%	88.56%	98.76%	84.15%	85.08%	79.67%	98.15%
85%	90%	0.8820	0.0104	0.8426	0.0139	96.93%	67.46%	89.62%	98.78%	87.29%	86.41%	81.02%	98.51%
90%	95%	0.8926	0.0097	0.8569	0.0130	97.20%	71.12%	89.72%	99.02%	89.38%	87.04%	82.21%	98.61%
95%	<	0.9065	0.0085	0.8753	0.0113	98.11%	74.49%	90.86%	99.12%	92.85%	88.44%	82.49%	99.16%
100%	100%	0.9288	0.0077	0.9050	0.0102	98.68%	83.87%	89.58%	99.38%	97.83%	88.63%	85.69%	99.41%

Region-based Baseline Classification. Monte Carlo Simulation using SVM-OAO
Setup 6 Train: [prop, 100%]
Test: [prop, 100%]

Threshold	OA		*Kappa*		User Accuracy				Producer Accuracy			
	Mean	**std**	**Mean**	**std**	**Forest**	**Crops**	**Bare Soil**	**Water**	**Forest**	**Crops**	**Bare Soil**	**Water**
50%	0.7719	0.0126	0.6959	0.0168	83.48%	48.97%	79.04%	97.26%	74.61%	65.68%	72.68%	92.81%
55%	0.7944	0.0127	0.7259	0.0169	85.64%	53.79%	80.62%	97.73%	76.86%	68.72%	75.38%	94.55%
60%	0.8102	0.0123	0.7469	0.0163	87.56%	56.43%	81.87%	98.20%	79.54%	71.25%	75.99%	95.38%
65%	0.8281	0.0117	0.7707	0.0156	89.46%	60.12%	83.03%	98.62%	82.08%	74.12%	77.41%	96.08%
70%	0.8430	0.0108	0.7907	0.0144	91.14%	63.20%	84.65%	98.22%	84.32%	76.40%	78.21%	97.29%
75%	0.8602	0.0120	0.8136	0.0161	92.44%	67.56%	85.46%	98.62%	86.20%	78.89%	80.54%	97.63%
80%	0.8754	0.0088	0.8339	0.0118	94.08%	70.78%	86.68%	98.64%	88.59%	81.40%	81.44%	98.23%
85%	0.8865	0.0100	0.8487	0.0134	95.00%	72.86%	87.97%	98.76%	89.94%	83.48%	82.39%	98.45%
90%	0.9000	0.0095	0.8667	0.0127	96.02%	76.67%	88.36%	98.96%	92.37%	84.97%	83.68%	98.73%
95%	0.9109	0.0082	0.8812	0.0109	97.09%	79.50%	88.65%	99.11%	95.05%	85.96%	84.14%	99.09%
100%	0.9288	0.0077	0.9050	0.0102	98.68%	83.87%	89.58%	99.38%	97.83%	88.63%	85.69%	99.41%

SOURCE: Author.

Table B.24 - Region-based Baseline Classification - Thematic Accuracy and Completeness for KNN-5 Setups 5 and 6.

Region-based Baseline Classification. Monte Carlo Simulation using KNN-5
Setup 5 — Train: [prop, prop + 5%[— Test: [prop, 100%]

Interval		OA		*Kappa*		User Accuracy				Producer Accuracy			
Inf	**Sup**	**Mean**	**std**	**Mean**	**std**	**Forest**	**Crops**	**Bare Soil**	**Water**	**Forest**	**Crops**	**Bare Soil**	**Water**
50%	55%	0.6249	0.0174	0.4998	0.0232	69.68%	32.57%	52.87%	94.83%	57.86%	37.25%	62.83%	87.83%
55%	60%	0.7047	0.0131	0.6063	0.0174	76.55%	35.66%	71.88%	97.79%	65.06%	49.88%	68.08%	92.90%
60%	65%	0.7465	0.0133	0.6620	0.0177	79.73%	43.45%	77.29%	98.13%	71.78%	57.98%	70.42%	94.14%
65%	70%	0.7852	0.0143	0.7136	0.0191	87.59%	48.42%	80.27%	97.82%	76.60%	66.13%	72.31%	96.44%
70%	75%	0.8093	0.0125	0.7457	0.0166	90.94%	53.96%	80.81%	98.00%	79.13%	71.40%	74.37%	97.20%
75%	80%	0.8390	0.0119	0.7853	0.0159	91.46%	62.05%	83.77%	98.32%	82.83%	75.63%	78.32%	97.76%
80%	85%	0.8571	0.0101	0.8095	0.0135	92.97%	68.60%	82.90%	98.37%	87.53%	76.46%	79.49%	98.58%
85%	90%	0.8789	0.0096	0.8385	0.0128	94.31%	73.02%	85.50%	98.72%	88.38%	80.89%	83.11%	98.59%
90%	95%	0.8943	0.0099	0.8590	0.0131	95.41%	76.36%	87.31%	98.63%	91.07%	83.66%	83.72%	98.96%
95%	< 100%	0.9025	0.0073	0.8700	0.0098	96.67%	77.40%	87.74%	99.20%	93.79%	84.48%	83.45%	99.03%
100%	100%	0.9325	0.0071	0.9100	0.0095	97.84%	86.25%	89.45%	99.48%	98.28%	88.03%	87.31%	99.46%

Region-based Baseline Classification. Monte Carlo Simulation using KNN-5
Setup 6 — Train: [prop, 100%] — Test: [prop, 100%]

Threshold	OA		*Kappa*		User Accuracy				Producer Accuracy			
	Mean	**std**	**Mean**	**std**	**Forest**	**Crops**	**Bare Soil**	**Water**	**Forest**	**Crops**	**Bare Soil**	**Water**
50%	0.7502	0.0151	0.6670	0.0201	77.97%	54.28%	72.07%	95.78%	74.15%	58.91%	71.99%	93.36%
55%	0.7761	0.0139	0.7014	0.0185	81.07%	58.08%	74.63%	96.65%	77.11%	62.38%	74.52%	95.12%
60%	0.7973	0.0128	0.7297	0.0171	83.63%	61.85%	76.22%	97.22%	80.28%	65.53%	76.12%	95.96%
65%	0.8157	0.0127	0.7543	0.0169	85.58%	64.76%	78.24%	97.70%	83.03%	68.31%	77.38%	96.71%
70%	0.8330	0.0101	0.7773	0.0135	87.70%	67.99%	79.91%	97.60%	85.17%	71.00%	78.77%	97.75%
75%	0.8521	0.0120	0.8028	0.0159	88.95%	71.86%	81.76%	98.27%	87.18%	73.81%	81.48%	98.01%
80%	0.8696	0.0120	0.8261	0.0159	91.12%	74.92%	83.39%	98.39%	89.92%	76.77%	82.42%	98.51%
85%	0.8832	0.0101	0.8443	0.0135	92.66%	77.14%	84.88%	98.62%	91.09%	79.33%	84.00%	98.62%
90%	0.8965	0.0095	0.8619	0.0127	93.93%	79.52%	86.30%	98.83%	93.25%	81.79%	84.57%	98.84%
95%	0.9080	0.0078	0.8773	0.0104	95.82%	81.35%	86.89%	99.14%	95.15%	83.56%	85.32%	99.05%
100%	0.9325	0.0071	0.9100	0.0095	97.84%	86.25%	89.45%	99.48%	98.28%	88.03%	87.31%	99.46%

SOURCE: Author.

B.2.3 Confusion matrices

The confusion matrices for Region-based Baseline Classification are presented in this appendix. Firstly the results of the SVM-OAO classifier is presented followed by KNN-5.

Table B.25 - Region-based Baseline Classification - Confusion Matrix for SVM-OAO Setup 1.

Region-based Baseline Classification. Monte Carlo Simulation using SVM-OAO
Setup 1 Train: [prop, prop + 5%[
Test: [50% 100%]

50-55%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**228.43**	2.6	4.64	14.33	250
Crops	121.44	**27.16**	100.79	0.61	250
Bare Soil	17.01	5.08	**212.74**	15.17	250
Water	1.98	0.12	0.86	**247.04**	250
Total	368.86	34.96	319.03	277.15	1000

55-60%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**226.78**	8.27	4.69	10.26	250
Crops	97.39	**50.8**	101.38	0.43	250
Bare Soil	14.12	7.48	**213.65**	14.75	250
Water	2.58	0.19	1.05	**246.18**	250
Total	340.87	66.74	320.77	271.62	1000

60-65%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**222.01**	14.04	5.16	8.79	250
Crops	80.32	**74.03**	95.38	0.27	250
Bare Soil	11.21	12.12	**214.72**	11.95	250
Water	3.14	0.37	1.68	**244.81**	250
Total	316.68	100.56	316.94	265.82	1000

65-70%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**221.01**	17.74	4.71	6.54	250
Crops	74.41	**90.77**	84.65	0.17	250
Bare Soil	9.57	19.75	**209.16**	11.52	250
Water	5.72	0.41	1.91	**241.96**	250
Total	310.71	128.67	300.43	260.19	1000

70-75%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**215.09**	23.98	4.42	6.51	250
Crops	64.93	**106.66**	78.26	0.15	250
Bare Soil	9.32	24.75	**205.43**	10.5	250
Water	5.88	0.4	2.49	**241.23**	250
Total	295.22	155.79	290.6	258.39	1000

75-80%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**209.4**	30.96	3.67	5.97	250
Crops	56.86	**123.47**	69.5	0.17	250
Bare Soil	7.73	32.78	**198.56**	10.93	250
Water	7.26	0.43	2.08	**240.23**	250
Total	281.25	187.64	273.81	257.3	1000

80-85%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**207.76**	33.94	3.28	5.02	250
Crops	54.19	**134.97**	60.73	0.11	250
Bare Soil	7.11	43.18	**191.36**	8.35	250
Water	9.05	0.48	4.6	**235.87**	250
Total	278.11	212.57	259.97	249.35	1000

85-90%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**200.33**	41.27	4.05	4.35	250
Crops	46.4	**146.4**	57.11	0.09	250
Bare Soil	5.53	48.31	**188.3**	7.86	250
Water	9.89	0.52	6.82	**232.77**	250
Total	262.15	236.5	256.28	245.07	1000

90-95%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**193**	48.98	3.91	4.11	250
Crops	40.21	**155.12**	54.61	0.06	250
Bare Soil	4.93	51.76	**185.64**	7.67	250
Water	10.08	0.61	8.38	**230.93**	250
Total	248.22	256.47	252.54	242.77	1000

95-100%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**182.35**	58.67	4.92	4.06	250
Crops	31.78	**164.77**	53.41	0.04	250
Bare Soil	2.93	54.92	**185.01**	7.14	250
Water	8.09	0.64	11.65	**229.62**	250
Total	225.15	279	254.99	240.86	1000

100%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**150.32**	90.75	5.46	3.47	250
Crops	17.25	**190.95**	41.78	0.02	250
Bare Soil	0.95	76.37	**165.24**	7.44	250
Water	5.88	0.87	16.43	**226.82**	250
Total	174.4	358.94	228.91	237.75	1000

SOURCE: Author.

Table B.26 - Region-based Baseline Classification - Confusion Matrix for KNN-5 Setup 1.

Region-based Baseline Classification. Monte Carlo Simulation using KNN-5

Setup 1 Train: [prop, prop + 5%[

Test: [50%, 100%]

50-55%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**176.67**	48.84	5.12	19.37	250
Grass	103.88	**85.22**	60.43	0.47	250
Bare Soil	18.54	83.12	**135.78**	12.56	250
Water	3.48	0.46	8.72	**237.34**	250
Total	302.57	217.64	210.05	269.74	1000

55-60%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**187.87**	46.65	4.58	10.9	250
Grass	80.12	**89.56**	80.03	0.29	250
Bare Soil	17.91	42.84	**177.13**	12.12	250
Water	3.41	0.59	3.26	**242.74**	250
Total	289.31	179.64	265	266.05	1000

60-65%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**190.1**	46.5	4.54	8.86	250
Grass	65.04	**108.63**	76.15	0.18	250
Bare Soil	13.14	38.47	**188.03**	10.36	250
Water	4.97	0.54	3.21	**241.28**	250
Total	273.25	194.14	271.93	260.68	1000

65-70%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**203.65**	34.38	6.21	5.76	250
Grass	64.08	**114.58**	71.2	0.14	250
Bare Soil	7.89	44.1	**189.6**	8.41	250
Water	7.81	0.5	5.9	**235.79**	250
Total	283.43	193.56	272.91	250.1	1000

70-75%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**205.89**	34.03	4.02	6.06	250
Grass	57.58	**125.13**	67.07	0.22	250
Bare Soil	11.01	43.56	**186.74**	8.69	250
Water	7.68	0.67	4.57	**237.08**	250
Total	282.16	203.39	262.4	252.05	1000

75-80%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**201.08**	39.35	4.95	4.62	250
Grass	50.49	**137.91**	61.36	0.24	250
Bare Soil	8.32	47.96	**185.25**	8.47	250
Water	12.46	0.52	5.13	**231.89**	250
Total	272.35	225.74	256.69	245.22	1000

80-85%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**192.75**	47.71	4.83	4.71	250
Grass	39.88	**155.02**	54.94	0.16	250
Bare Soil	5.6	58.05	**180.19**	6.16	250
Water	10.62	0.52	8.5	**230.36**	250
Total	248.85	261.3	248.46	241.39	1000

85-90%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**193.46**	45.93	5.6	5.01	250
Grass	40.84	**159.52**	49.51	0.13	250
Bare Soil	5.41	59.66	**178.03**	6.9	250
Water	9.09	0.53	7.71	**232.67**	250
Total	248.8	265.64	240.85	244.71	1000

90-95%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**181.77**	58.92	5.79	3.52	250
Grass	32.9	**169.01**	48.05	0.04	250
Bare Soil	4.56	62.04	**176.93**	6.47	250
Water	16.47	0.66	9.52	**223.35**	250
Total	235.7	290.63	240.29	233.38	1000

95-100%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**172.42**	66.93	6.65	4	250
Grass	27.5	**174.88**	47.55	0.07	250
Bare Soil	2.95	63.95	**176.33**	6.77	250
Water	8.3	0.74	11.85	**229.11**	250
Total	211.17	306.5	242.38	239.95	1000

100%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**132.91**	104.98	7.21	4.9	250
Grass	12.29	**198.79**	38.82	0.1	250
Bare Soil	0.86	78.9	**163.53**	6.71	250
Water	4.37	1.28	12.13	**232.22**	250
Total	150.43	383.95	221.69	243.93	1000

SOURCE: Author.

Table B.27 - Region-based Baseline Classification - Confusion Matrix for SVM-OAO Setup 2.

Region-based Baseline Classification. Monte Carlo Simulation using SVM-OAO
Setup 2 Train: [prop, prop + 5%[
Test: [100%]

50-55%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**247.5**	0	0.16	2.34	250
Crops	100.86	**45.17**	103.66	0.31	250
Bare Soil	6.23	2.47	**235.83**	5.47	250
Water	0	0	0.27	**249.73**	250
Total	354.59	47.64	339.92	257.85	1000

55-60%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**249.72**	0	0.28	0	250
Crops	71.94	**69.55**	108.38	0.13	250
Bare Soil	5.8	1.8	**237.52**	4.88	250
Water	0	0	0.31	**249.69**	250
Total	327.46	71.35	346.49	254.7	1000

60-65%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**249.45**	0	0.55	0	250
Crops	52.81	**100.24**	96.95	0	250
Bare Soil	3.68	1.77	**241.21**	3.34	250
Water	0	0	0.35	**249.65**	250
Total	305.94	102.01	339.06	252.99	1000

65-70%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**249.37**	0	0.63	0	250
Crops	48.06	**116.36**	85.58	0	250
Bare Soil	2.74	2.75	**241.18**	3.33	250
Water	0	0	0.35	**249.65**	250
Total	300.17	119.11	327.74	252.98	1000

70-75%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**249.37**	0	0.63	0	250
Crops	36.84	**135.4**	77.76	0	250
Bare Soil	2.83	4.21	**240.01**	2.95	250
Water	0	0	0.35	**249.65**	250
Total	289.04	139.61	318.75	252.6	1000

75-80%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**249.18**	0.19	0.63	0	250
Crops	31.32	**148.71**	69.97	0	250
Bare Soil	2.66	5.49	**238.61**	3.24	250
Water	0	0	0.35	**249.65**	250
Total	283.16	154.39	309.56	252.89	1000

80-85%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**249.15**	0.22	0.63	0	250
Crops	29.76	**163.54**	56.7	0	250
Bare Soil	2.34	10.81	**235.11**	1.74	250
Water	0	0	0.73	**249.27**	250
Total	281.25	174.57	293.17	251.01	1000

85-90%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**248.96**	0.41	0.63	0	250
Crops	24.34	**172.08**	53.58	0	250
Bare Soil	1.04	11.54	**236.18**	1.24	250
Water	0	0	0.93	**249.07**	250
Total	274.34	184.03	291.32	250.31	1000

90-95%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**248.82**	0.55	0.63	0	250
Crops	21.79	**178.72**	49.49	0	250
Bare Soil	1.05	13.51	**234**	1.44	250
Water	0.03	0	0.98	**248.99**	250
Total	271.69	192.78	285.1	250.43	1000

95-100%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**248.82**	0.55	0.63	0	250
Crops	15.11	**185.47**	49.42	0	250
Bare Soil	0.69	13.6	**234.76**	0.95	250
Water	0	0	1.09	**248.91**	250
Total	264.62	199.62	285.9	249.86	1000

100%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**247.52**	1.9	0.58	0	250
Crops	6.88	**204.43**	38.69	0	250
Bare Soil	0.46	21.73	**226.28**	1.53	250
Water	0	0	1.3	**248.7**	250
Total	254.86	228.06	266.85	250.23	1000

SOURCE: Author.

Table B.28 - Region-based Baseline Classification - Confusion Matrix for KNN-5 Setup 2.

Region-based Baseline Classification. Monte Carlo Simulation using KNN-5
Setup 2 Train: [prop, prop + 5%[
Test: [100%]

50-55%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**199.26**	28.3	0.5	21.94	250
Grass	90.5	**92.94**	66.43	0.13	250
Bare Soil	6.95	93.53	**144.66**	4.86	250
Water	0.03	0	1.71	**248.26**	250
Total	296.74	214.77	213.3	275.19	1000

55-60%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**227.6**	20.78	0.05	1.57	250
Grass	70.68	**90.54**	88.76	0.02	250
Bare Soil	8.26	36.72	**201.78**	3.24	250
Water	0	0	0.8	**249.2**	250
Total	306.54	148.04	291.39	254.03	1000

60-65%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**239.24**	8.53	0.09	2.14	250
Grass	49.37	**124.2**	76.43	0	250
Bare Soil	5.34	23.33	**218.59**	2.74	250
Water	0	0	0.79	**249.21**	250
Total	293.95	156.06	295.9	254.09	1000

65-70%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**245.32**	4.14	0.54	0	250
Grass	43.98	**135.61**	70.41	0	250
Bare Soil	1.59	23.87	**222.61**	1.93	250
Water	0.04	0	0.8	**249.16**	250
Total	290.93	163.62	294.36	251.09	1000

70-75%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**247.79**	2.05	0.16	0	250
Grass	33.64	**151.68**	64.66	0.02	250
Bare Soil	3.81	28.39	**216.18**	1.62	250
Water	0.09	0	0.76	**249.15**	250
Total	285.33	182.12	281.76	250.79	1000

75-80%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**246.88**	2.57	0.55	0	250
Grass	28.43	**162.13**	59.44	0	250
Bare Soil	3.21	14.65	**230.33**	1.81	250
Water	0.05	0	0.45	**249.5**	250
Total	278.57	179.35	290.77	251.31	1000

80-85%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**247.72**	2	0.28	0	250
Grass	21.58	**176.8**	51.62	0	250
Bare Soil	1.41	23.41	**224.12**	1.06	250
Water	0.02	0	1.06	**248.92**	250
Total	270.73	202.21	277.08	249.98	1000

85-90%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**248.04**	1.6	0.36	0	250
Grass	19.64	**187.97**	42.39	0	250
Bare Soil	1.96	18.79	**228.03**	1.22	250
Water	0.02	0	0.95	**249.03**	250
Total	269.66	208.36	271.73	250.25	1000

90-95%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**247.85**	1.53	0.62	0	250
Grass	15.26	**194.43**	40.31	0	250
Bare Soil	1.2	22.2	**225.67**	0.93	250
Water	0.42	0	1.18	**248.4**	250
Total	264.73	218.16	267.78	249.33	1000

95-100%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**247.99**	1.38	0.63	0	250
Grass	12.25	**194.68**	43.07	0	250
Bare Soil	0.97	21.06	**226.71**	1.26	250
Water	0	0	0.83	**249.17**	250
Total	261.21	217.12	271.24	250.43	1000

100%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**246.21**	3.41	0.38	0	250
Grass	3.36	**222.95**	23.69	0	250
Bare Soil	0.32	17.85	**230.55**	1.28	250
Water	0	0.01	1.12	**248.87**	250
Total	249.89	244.22	255.74	250.15	1000

SOURCE: Author.

Table B.29 - Region-based Baseline Classification - Confusion Matrix for SVM-OAO Setup 3.

Region-based Baseline Classification. Monte Carlo Simulation using SVM-OAO
Setup 3 Train: [prop, 100%]
Test: [50% 100%]

50% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**208.42**	30.26	3.53	7.79	250
Crops	58.23	**122.81**	68.79	0.17	250
Bare Soil	7.91	33.64	**197.49**	10.96	250
Water	4.21	0.44	2.16	**243.19**	250
Total	278.77	187.15	271.97	262.11	1000

55% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**205.86**	33.24	3.67	7.23	250
Crops	54.49	**127.93**	67.41	0.17	250
Bare Soil	7.29	35.68	**196.52**	10.51	250
Water	4.81	0.46	2.43	**242.3**	250
Total	272.45	197.31	270.03	260.21	1000

60% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**203.62**	36.01	3.67	6.7	250
Crops	51.7	**134.15**	64	0.15	250
Bare Soil	6.48	39.76	**193.9**	9.86	250
Water	5.33	0.47	2.93	**241.27**	250
Total	267.13	210.39	264.5	257.98	1000

65% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**202.05**	38.38	3.56	6.01	250
Crops	49.42	**139.57**	60.88	0.13	250
Bare Soil	6.11	43.7	**190.67**	9.52	250
Water	6.35	0.49	3.4	**239.76**	250
Total	263.93	222.14	258.51	255.42	1000

70% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**199.03**	41.66	3.45	5.86	250
Crops	45.94	**144.91**	59.03	0.12	250
Bare Soil	5.78	46.42	**188.82**	8.98	250
Water	6.45	0.49	4.31	**238.75**	250
Total	257.2	233.48	255.61	253.71	1000

75% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**195.41**	45.54	3.73	5.32	250
Crops	42.53	**150.98**	56.38	0.11	250
Bare Soil	5.23	49.24	**186.82**	8.71	250
Water	7.58	0.57	4.69	**237.16**	250
Total	250.75	246.33	251.62	251.3	1000

80% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**192.47**	48.89	3.81	4.83	250
Crops	39.69	**156.75**	53.47	0.09	250
Bare Soil	4.69	53.68	**183.4**	8.23	250
Water	8.07	0.56	6.67	**234.7**	250
Total	244.92	259.88	247.35	247.85	1000

85% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**187.03**	54.35	4.38	4.24	250
Crops	35.4	**162.88**	51.67	0.05	250
Bare Soil	3.76	57.54	**181.04**	7.66	250
Water	8.36	0.62	9.28	**231.74**	250
Total	234.55	275.39	246.37	243.69	1000

90% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**180.91**	60.55	4.44	4.1	250
Crops	30.98	**169.51**	49.46	0.05	250
Bare Soil	3.26	59.83	**179.14**	7.77	250
Water	8.2	0.65	10.25	**230.9**	250
Total	223.35	290.54	243.29	242.82	1000

95% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**170.7**	70.29	5.05	3.96	250
Crops	25.42	**177.39**	47.17	0.02	250
Bare Soil	1.9	64.4	**176.35**	7.35	250
Water	6.63	0.73	13.4	**229.24**	250
Total	204.65	312.81	241.97	240.57	1000

100% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**150.32**	90.75	5.46	3.47	250
Crops	17.25	**190.95**	41.78	0.02	250
Bare Soil	0.95	76.37	**165.24**	7.44	250
Water	5.88	0.87	16.43	**226.82**	250
Total	174.4	358.94	228.91	237.75	1000

SOURCE: Author.

Table B.30 - Region-based Baseline Classification - Confusion Matrix for KNN-5 Setup 3.

Region-based Baseline Classification. Monte Carlo Simulation using KNN-5
Setup 3 Train: [prop, 100%]
Test: [50%, 100%]

50% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**195.76**	41.69	4.83	7.72	250
Grass	52.18	**138.6**	59.04	0.18	250
Bare Soil	8.87	50.3	**181.24**	9.59	250
Water	5.14	0.52	4.73	**239.61**	250
Total	261.95	231.11	249.84	257.1	1000

55% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**194.69**	43.68	4.88	6.75	250
Grass	47.7	**142.94**	59.19	0.17	250
Bare Soil	7.48	51.15	**182.61**	8.76	250
Water	5.91	0.63	5.01	**238.45**	250
Total	255.78	238.4	251.69	254.13	1000

60% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**193.34**	45.5	5.2	5.96	250
Grass	44.79	**148.81**	56.26	0.14	250
Bare Soil	6.34	54.51	**180.51**	8.64	250
Water	7.12	0.61	5.38	**236.89**	250
Total	251.59	249.43	247.35	251.63	1000

65% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**192.25**	47.12	5.01	5.62	250
Grass	43.25	**152.17**	54.45	0.13	250
Bare Soil	5.98	56.06	**180.14**	7.82	250
Water	8.4	0.61	5.95	**235.04**	250
Total	249.88	255.96	245.55	248.61	1000

70% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**188.99**	50.33	5.37	5.31	250
Grass	39.59	**158.71**	51.56	0.14	250
Bare Soil	5.99	58.36	**178.23**	7.42	250
Water	8.37	0.63	7.05	**233.95**	250
Total	242.94	268.03	242.21	246.82	1000

75% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**185.7**	53.98	5.18	5.14	250
Grass	36.1	**163.86**	49.94	0.1	250
Bare Soil	5.07	60.06	**177.45**	7.42	250
Water	9.41	0.58	6.46	**233.55**	250
Total	236.28	278.48	239.03	246.21	1000

80% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**183.13**	56.54	5.48	4.85	250
Grass	33.4	**168.47**	48.03	0.1	250
Bare Soil	4.31	62.96	**175.86**	6.87	250
Water	9.44	0.7	8.98	**230.88**	250
Total	230.28	288.67	238.35	242.7	1000

85% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**178**	61.33	6.32	4.35	250
Grass	30.63	**172.87**	46.43	0.07	250
Bare Soil	3.46	64.99	**174.65**	6.9	250
Water	9.84	0.69	10.23	**229.24**	250
Total	221.93	299.88	237.63	240.56	1000

90% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**170.44**	68.98	6.46	4.12	250
Grass	26.73	**178.43**	44.78	0.06	250
Bare Soil	2.94	66.66	**173.54**	6.86	250
Water	9.81	0.71	11.59	**227.89**	250
Total	209.92	314.78	236.37	238.93	1000

95% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**160.77**	77.78	7.16	4.29	250
Grass	22.12	**184.53**	43.25	0.1	250
Bare Soil	2.06	70.03	**171.14**	6.77	250
Water	6.37	0.86	12.56	**230.21**	250
Total	191.32	333.2	234.11	241.37	1000

100% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**132.91**	104.98	7.21	4.9	250
Grass	12.29	**198.79**	38.82	0.1	250
Bare Soil	0.86	78.9	**163.53**	6.71	250
Water	4.37	1.28	12.13	**232.22**	250
Total	150.43	383.95	221.69	243.93	1000

SOURCE: Author.

Table B.31 - Region-based Baseline Classification - Confusion Matrix for SVM-OAO Setup 4.

Region-based Baseline Classification. Monte Carlo Simulation using SVM-OAO

Setup 4 Train: [prop, 100%]

Test: [100%]

50%

MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**249.2**	0.17	0.63	0	250
Crops	31.02	**150.85**	68.13	0	250
Bare Soil	2.45	6.23	**238.32**	3	250
Water	0	0	0.35	**249.65**	250
Total	282.67	157.25	307.43	252.65	1000

55%

MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**249.12**	0.25	0.63	0	250
Crops	29.59	**155.65**	64.76	0	250
Bare Soil	2.2	6.9	**238.06**	2.84	250
Water	0	0	0.37	**249.63**	250
Total	280.91	162.8	303.82	252.47	1000

60%

MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**249.12**	0.25	0.63	0	250
Crops	29.59	**155.65**	64.76	0	250
Bare Soil	2.2	6.9	**238.06**	2.84	250
Water	0	0	0.37	**249.63**	250
Total	280.91	162.8	303.82	252.47	1000

65%

MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**249.09**	0.28	0.63	0	250
Crops	27.6	**160.81**	61.59	0	250
Bare Soil	1.93	8.35	**237.18**	2.54	250
Water	0	0	0.51	**249.49**	250
Total	278.62	169.44	299.91	252.03	1000

70%

MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**249.01**	0.36	0.63	0	250
Crops	25.82	**167.29**	56.89	0	250
Bare Soil	1.7	10.27	**235.76**	2.27	250
Water	0	0	0.55	**249.45**	250
Total	276.53	177.92	293.83	251.72	1000

75%

MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**248.9**	0.47	0.63	0	250
Crops	24.28	**171.77**	53.95	0	250
Bare Soil	1.33	11.67	**235.25**	1.75	250
Water	0	0	0.75	**249.25**	250
Total	274.51	183.91	290.58	251	1000

80%

MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**248.81**	0.56	0.63	0	250
Crops	22.35	**176.13**	51.52	0	250
Bare Soil	1.05	12.58	**235.04**	1.33	250
Water	0	0	0.89	**249.11**	250
Total	272.21	189.27	288.08	250.44	1000

85%

MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**248.82**	0.55	0.63	0	250
Crops	19.73	**180.74**	49.53	0	250
Bare Soil	0.78	13.39	**234.51**	1.32	250
Water	0.02	0	0.96	**249.02**	250
Total	269.35	194.68	285.63	250.34	1000

90%

MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**248.82**	0.55	0.63	0	250
Crops	17.25	**185.74**	47.01	0	250
Bare Soil	0.89	14.96	**232.86**	1.29	250
Water	0.05	0	0.96	**248.99**	250
Total	267.01	201.25	281.46	250.28	1000

95%

MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**248.78**	0.59	0.63	0	250
Crops	12.22	**191.81**	45.97	0	250
Bare Soil	0.69	15.18	**233.1**	1.03	250
Water	0	0	1.16	**248.84**	250
Total	261.69	207.58	280.86	249.87	1000

100%

MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**247.52**	1.9	0.58	0	250
Crops	6.88	**204.43**	38.69	0	250
Bare Soil	0.46	21.73	**226.28**	1.53	250
Water	0	0	1.3	**248.7**	250
Total	254.86	228.06	266.85	250.23	1000

SOURCE: Author.

Table B.32 - Region-based Baseline Classification - Confusion Matrix for KNN-5 Setup 4.

Region-based Baseline Classification. Monte Carlo Simulation using KNN-5

Setup 4 Train: [prop, 100%]

Test: [100%]

50% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**247.61**	1.9	0.42	0.07	250
Grass	32.29	**161.56**	56.15	0	250
Bare Soil	3.07	20.76	**223.77**	2.4	250
Water	0.03	0	0.77	**249.2**	250
Total	283	184.22	281.11	251.67	1000

55% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**247.62**	1.96	0.41	0.01	250
Grass	28.55	**165.95**	55.5	0	250
Bare Soil	2.68	18.04	**227.18**	2.1	250
Water	0.01	0	0.77	**249.22**	250
Total	278.86	185.95	283.86	251.33	1000

60% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**247.68**	1.83	0.45	0.04	250
Grass	24.9	**172.63**	52.47	0	250
Bare Soil	2.07	17.96	**228.15**	1.82	250
Water	0.03	0	0.76	**249.21**	250
Total	274.68	192.42	281.83	251.07	1000

65% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**247.78**	1.76	0.46	0	250
Grass	22.56	**177.73**	49.71	0	250
Bare Soil	1.87	19.37	**227.25**	1.51	250
Water	0.05	0	0.91	**249.04**	250
Total	272.26	198.86	278.33	250.55	1000

70% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**247.86**	1.65	0.49	0	250
Grass	20.22	**183.11**	46.66	0.01	250
Bare Soil	1.59	19.45	**227.62**	1.34	250
Water	0.03	0	0.83	**249.14**	250
Total	269.7	204.21	275.6	250.49	1000

75% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**247.93**	1.49	0.58	0	250
Grass	18.15	**188.29**	43.56	0	250
Bare Soil	1.37	19.6	**227.85**	1.18	250
Water	0.02	0	0.88	**249.1**	250
Total	267.47	209.38	272.87	250.28	1000

80% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**248**	1.46	0.54	0	250
Grass	15.33	**192.85**	41.82	0	250
Bare Soil	1.1	18.99	**228.72**	1.19	250
Water	0.08	0	1.03	**248.89**	250
Total	264.51	213.3	272.11	250.08	1000

85% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**248.19**	1.2	0.61	0	250
Grass	14.49	**197.41**	38.1	0	250
Bare Soil	0.78	19.46	**228.56**	1.2	250
Water	0.14	0	1.05	**248.81**	250
Total	263.6	218.07	268.32	250.01	1000

90% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**248.12**	1.24	0.64	0	250
Grass	12.04	**201.38**	36.58	0	250
Bare Soil	0.84	20.03	**228.09**	1.04	250
Water	0.09	0	1.1	**248.81**	250
Total	261.09	222.65	266.41	249.85	1000

95% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**247.99**	1.38	0.63	0	250
Grass	9.54	**205.96**	34.5	0	250
Bare Soil	0.79	16.91	**231.23**	1.07	250
Water	0	0	0.92	**249.08**	250
Total	258.32	224.25	267.28	250.15	1000

100% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**246.21**	3.41	0.38	0	250
Grass	3.36	**222.95**	23.69	0	250
Bare Soil	0.32	17.85	**230.55**	1.28	250
Water	0	0.01	1.12	**248.87**	250
Total	249.89	244.22	255.74	250.15	1000

SOURCE: Author.

Table B.33 - Region-based Baseline Classification - Confusion Matrix for SVM-OAO Setup 5.

Region-based Baseline Classification. Monte Carlo Simulation using SVM-OAO
Setup 5 Train: [prop, prop + 5%[
Test: [prop, 100%]

50-55%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**249.2**	0.17	0.63	0	250
Crops	31.02	**150.85**	68.13	0	250
Bare Soil	2.45	6.23	**238.32**	3	250
Water	0	0	0.35	**249.65**	250
Total	282.67	157.25	307.43	252.65	1000

55-60%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**249.12**	0.25	0.63	0	250
Crops	29.59	**155.65**	64.76	0	250
Bare Soil	2.2	6.9	**238.06**	2.84	250
Water	0	0	0.37	**249.63**	250
Total	280.91	162.8	303.82	252.47	1000

60-65%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**249.12**	0.25	0.63	0	250
Crops	29.59	**155.65**	64.76	0	250
Bare Soil	2.2	6.9	**238.06**	2.84	250
Water	0	0	0.37	**249.63**	250
Total	280.91	162.8	303.82	252.47	1000

65-70%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**249.09**	0.28	0.63	0	250
Crops	27.6	**160.81**	61.59	0	250
Bare Soil	1.93	8.35	**237.18**	2.54	250
Water	0	0	0.51	**249.49**	250
Total	278.62	169.44	299.91	252.03	1000

70-75%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**249.01**	0.36	0.63	0	250
Crops	25.82	**167.29**	56.89	0	250
Bare Soil	1.7	10.27	**235.76**	2.27	250
Water	0	0	0.55	**249.45**	250
Total	276.53	177.92	293.83	251.72	1000

75-80%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**248.9**	0.47	0.63	0	250
Crops	24.28	**171.77**	53.95	0	250
Bare Soil	1.33	11.67	**235.25**	1.75	250
Water	0	0	0.75	**249.25**	250
Total	274.51	183.91	290.58	251	1000

80-85%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**248.81**	0.56	0.63	0	250
Crops	22.35	**176.13**	51.52	0	250
Bare Soil	1.05	12.58	**235.04**	1.33	250
Water	0	0	0.89	**249.11**	250
Total	272.21	189.27	288.08	250.44	1000

85-90%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**248.82**	0.55	0.63	0	250
Crops	19.73	**180.74**	49.53	0	250
Bare Soil	0.78	13.39	**234.51**	1.32	250
Water	0.02	0	0.96	**249.02**	250
Total	269.35	194.68	285.63	250.34	1000

90-95%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**248.82**	0.55	0.63	0	250
Crops	17.25	**185.74**	47.01	0	250
Bare Soil	0.89	14.96	**232.86**	1.29	250
Water	0.05	0	0.96	**248.99**	250
Total	267.01	201.25	281.46	250.28	1000

95-100%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**248.78**	0.59	0.63	0	250
Crops	12.22	**191.81**	45.97	0	250
Bare Soil	0.69	15.18	**233.1**	1.03	250
Water	0	0	1.16	**248.84**	250
Total	261.69	207.58	280.86	249.87	1000

100%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Crops**	**Bare Soil**	**Water**	**Total**
Forest	**247.52**	1.9	0.58	0	250
Crops	6.88	**204.43**	38.69	0	250
Bare Soil	0.46	21.73	**226.28**	1.53	250
Water	0	0	1.3	**248.7**	250
Total	254.86	228.06	266.85	250.23	1000

SOURCE: Author.

Table B.34 - Region-based Baseline Classification - Confusion Matrix for KNN-5 Setup 5.

Region-based Baseline Classification. Monte Carlo Simulation using KNN-5

Setup 5 Train: [prop, prop + 5%[

Test: [prop, 100%]

50-55%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**174.21**	50.78	5.62	19.39	250
Grass	104.53	**81.42**	63.6	0.45	250
Bare Soil	18.93	85.88	**132.17**	13.02	250
Water	3.43	0.51	8.98	**237.08**	250
Total	301.1	218.59	210.37	269.94	1000

55-60%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**191.37**	45.82	3.67	9.14	250
Grass	82.48	**89.14**	78.17	0.21	250
Bare Soil	17.64	43.32	**179.71**	9.33	250
Water	2.65	0.44	2.43	**244.48**	250
Total	294.14	178.72	263.98	263.16	1000

60-65%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**199.33**	40.75	2.93	6.99	250
Grass	65.12	**108.62**	76.15	0.11	250
Bare Soil	11.07	37.55	**193.22**	8.16	250
Water	2.18	0.42	2.07	**245.33**	250
Total	277.7	187.34	274.37	260.59	1000

65-70%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**218.97**	24.47	3.49	3.07	250
Grass	58.62	**121.04**	70.27	0.07	250
Bare Soil	6.36	37.09	**200.67**	5.88	250
Water	1.92	0.43	3.1	**244.55**	250
Total	285.87	183.03	277.53	253.57	1000

70-75%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**227.36**	18.59	1.85	2.2	250
Grass	49.26	**134.9**	65.77	0.07	250
Bare Soil	8.36	34.82	**202.02**	4.8	250
Water	2.35	0.63	2.01	**245.01**	250
Total	287.33	188.94	271.65	252.08	1000

75-80%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**228.66**	18.72	1.37	1.25	250
Grass	39.83	**155.12**	54.94	0.11	250
Bare Soil	5.25	31.04	**209.43**	4.28	250
Water	2.31	0.22	1.67	**245.8**	250
Total	276.05	205.1	267.41	251.44	1000

80-85%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**232.43**	15.31	1.34	0.92	250
Grass	28.51	**171.51**	49.95	0.03	250
Bare Soil	2.94	37.23	**207.25**	2.58	250
Water	1.65	0.26	2.17	**245.92**	250
Total	265.53	224.31	260.71	249.45	1000

85-90%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**235.78**	12.11	1.38	0.73	250
Grass	26.91	**182.55**	40.52	0.02	250
Bare Soil	2.77	30.69	**213.76**	2.78	250
Water	1.32	0.34	1.54	**246.8**	250
Total	266.78	225.69	257.2	250.33	1000

90-95%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**238.53**	9.51	1.38	0.58	250
Grass	19.92	**190.9**	39.17	0.01	250
Bare Soil	2.2	27.51	**218.28**	2.01	250
Water	1.26	0.26	1.91	**246.57**	250
Total	261.91	228.18	260.74	249.17	1000

95-100%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**241.68**	7.28	0.84	0.2	250
Grass	15.26	**193.5**	41.2	0.04	250
Bare Soil	0.73	27.73	**219.34**	2.2	250
Water	0.02	0.53	1.46	**247.99**	250
Total	257.69	229.04	262.84	250.43	1000

100%

MEAN CONFUSION MATRIX					
CLASS	**Forest**	**Grass**	**Bare Soil**	**Water**	**Total**
Forest	**244.59**	4.76	0.65	0	250
Grass	3.84	**215.62**	30.54	0	250
Bare Soil	0.45	24.57	**223.62**	1.36	250
Water	0	0	1.31	**248.69**	250
Total	248.88	244.95	256.12	250.05	1000

SOURCE: Author.

Table B.35 - Region-based Baseline Classification - Confusion Matrix for SVM-OAO Setup 6.

Region-based Baseline Classification. Monte Carlo Simulation using SVM-OAO
Setup 6 Train: [prop, 100%]
Test: [prop, 100%]

50% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**208.7**	29.96	3.61	7.73	250
Crops	58.76	**122.43**	68.61	0.2	250
Bare Soil	7.92	33.57	**197.61**	10.9	250
Water	4.33	0.45	2.06	**243.16**	250
Total	279.71	186.41	271.89	261.99	1000

55% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**214.1**	27.79	2.6	5.51	250
Crops	54.05	**134.47**	61.35	0.13	250
Bare Soil	6.96	33.05	**201.54**	8.45	250
Water	3.44	0.37	1.87	**244.32**	250
Total	278.55	195.68	267.36	258.41	1000

60% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**218.91**	24.7	2.04	4.35	250
Crops	48.03	**141.08**	60.8	0.09	250
Bare Soil	5.92	31.94	**204.68**	7.46	250
Water	2.37	0.28	1.84	**245.51**	250
Total	275.23	198	269.36	257.41	1000

65% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**223.64**	21.18	1.95	3.23	250
Crops	42.71	**150.31**	56.91	0.07	250
Bare Soil	4.72	30.95	**207.57**	6.76	250
Water	1.39	0.36	1.71	**246.54**	250
Total	272.46	202.8	268.14	256.6	1000

70% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**227.85**	18.29	1.86	2	250
Crops	36.95	**158.01**	55.01	0.03	250
Bare Soil	3.46	30.1	**211.62**	4.82	250
Water	1.96	0.41	2.08	**245.55**	250
Total	270.22	206.81	270.57	252.4	1000

75% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**231.11**	16.06	1.45	1.38	250
Crops	32.54	**168.89**	48.54	0.03	250
Bare Soil	2.86	28.93	**213.64**	4.57	250
Water	1.6	0.21	1.63	**246.56**	250
Total	268.11	214.09	265.26	252.54	1000

80% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**235.21**	12.68	1.26	0.85	250
Crops	26.77	**176.94**	46.29	0	250
Bare Soil	2.26	27.45	**216.7**	3.59	250
Water	1.26	0.3	1.85	**246.59**	250
Total	265.5	217.37	266.1	251.03	1000

85% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**237.51**	10.42	1.42	0.65	250
Crops	23.74	**182.15**	44.1	0.01	250
Bare Soil	1.67	25.17	**219.93**	3.23	250
Water	1.15	0.46	1.49	**246.9**	250
Total	264.07	218.2	266.94	250.79	1000

90% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**240.04**	8.09	1.23	0.64	250
Crops	18.25	**191.67**	40.08	0	250
Bare Soil	0.97	25.59	**220.89**	2.55	250
Water	0.6	0.23	1.76	**247.41**	250
Total	259.86	225.58	263.96	250.6	1000

95% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**242.73**	6.4	0.71	0.16	250
Crops	11.96	**198.74**	39.3	0	250
Bare Soil	0.65	25.62	**221.62**	2.11	250
Water	0.02	0.43	1.77	**247.78**	250
Total	255.36	231.19	263.4	250.05	1000

100% — MEAN CONFUSION MATRIX

CLASS	Forest	Crops	Bare Soil	Water	Total
Forest	**246.7**	2.67	0.63	0	250
Crops	5.12	**209.67**	35.21	0	250
Bare Soil	0.35	24.22	**223.95**	1.48	250
Water	0	0	1.56	**248.44**	250
Total	252.17	236.56	261.35	249.92	1000

SOURCE: Author.

Table B.36 - Region-based Baseline Classification - Confusion Matrix for KNN-5 Setup 6.

Region-based Baseline Classification. Monte Carlo Simulation using KNN-5

Setup 6 Train: [prop, 100%]

Test: [prop, 100%]

50% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**194.92**	42.32	5.22	7.54	250
Grass	54.09	**135.7**	60.01	0.2	250
Bare Soil	8.75	51.78	**180.17**	9.3	250
Water	5.12	0.56	4.87	**239.45**	250
Total	262.88	230.36	250.27	256.49	1000

55% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**202.67**	38.27	4.02	5.04	250
Grass	48.57	**145.21**	56.06	0.16	250
Bare Soil	7.43	48.8	**186.57**	7.2	250
Water	4.15	0.51	3.71	**241.63**	250
Total	262.82	232.79	250.36	254.03	1000

60% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**209.07**	33.67	3.41	3.85	250
Grass	42.34	**154.62**	52.94	0.1	250
Bare Soil	5.88	47.28	**190.55**	6.29	250
Water	3.12	0.4	3.42	**243.06**	250
Total	260.41	235.97	250.32	253.3	1000

65% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**213.95**	30.13	2.97	2.95	250
Grass	37.22	**161.91**	50.83	0.04	250
Bare Soil	4.55	44.54	**195.6**	5.31	250
Water	1.95	0.43	3.38	**244.24**	250
Total	257.67	237.01	252.78	252.54	1000

70% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**219.24**	26.3	2.62	1.84	250
Grass	31.85	**169.97**	48.15	0.03	250
Bare Soil	3.86	42.63	**199.77**	3.74	250
Water	2.45	0.48	3.08	**243.99**	250
Total	257.4	239.38	253.62	249.6	1000

75% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**222.37**	24.16	2.05	1.42	250
Grass	28.03	**179.64**	42.29	0.04	250
Bare Soil	2.77	39.3	**204.39**	3.54	250
Water	1.9	0.29	2.13	**245.68**	250
Total	255.07	243.39	250.86	250.68	1000

80% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**227.81**	19.54	1.72	0.93	250
Grass	22.09	**187.3**	40.58	0.03	250
Bare Soil	1.99	36.77	**208.47**	2.77	250
Water	1.47	0.37	2.18	**245.98**	250
Total	253.36	243.98	252.95	249.71	1000

85% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**231.64**	16.13	1.57	0.66	250
Grass	19.98	**192.84**	37.16	0.02	250
Bare Soil	1.43	33.6	**212.21**	2.76	250
Water	1.26	0.51	1.68	**246.55**	250
Total	254.31	243.08	252.62	249.99	1000

90% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**234.83**	12.94	1.56	0.67	250
Grass	15.34	**198.8**	35.82	0.04	250
Bare Soil	1.06	30.99	**215.75**	2.2	250
Water	0.6	0.34	1.98	**247.08**	250
Total	251.83	243.07	255.11	249.99	1000

95% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**239.55**	9.46	0.8	0.19	250
Grass	11.58	**203.38**	34.98	0.06	250
Bare Soil	0.59	30.05	**217.23**	2.13	250
Water	0.05	0.51	1.59	**247.85**	250
Total	251.77	243.4	254.6	250.23	1000

100% — MEAN CONFUSION MATRIX

CLASS	Forest	Grass	Bare Soil	Water	Total
Forest	**244.59**	4.76	0.65	0	250
Grass	3.84	**215.62**	30.54	0	250
Bare Soil	0.45	24.57	**223.62**	1.36	250
Water	0	0	1.31	**248.69**	250
Total	248.88	244.95	256.12	250.05	1000

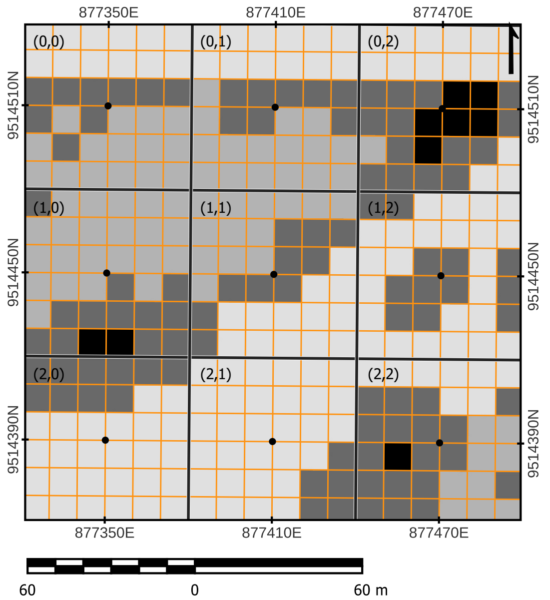
SOURCE: Author.

APPENDIX C - IMAGES IN BLACK AND WHITE (B&W)

In order to increase accessibility, some images are presented here in Black and White. They are all mentioned on the text, to be accessed as needed.

Figure C.1 - RSS Part I running the grid mask through the LR image in Black and White.

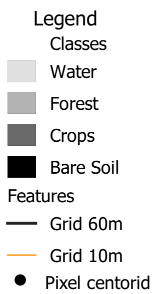
(a) Thematic Map



(b) Geodataframe

Pixel	Modal Class	prop	Centroid (E,N)	Candidade Sample
(0,0)	Forest	41.67%	(877350, 9514510)	No
(0,1)	Forest	44.44%	(877410, 9514510)	No
(0,2)	Water	41.47%	(877470, 9514510)	No
(1,0)	Forest	61.11%	(877350, 9514450)	Yes
(1,1)	Water*	36.11%	(877410, 9514450)	No
(1,2)	Water	63.89%	(877470, 9514450)	Yes
(2,0)	Water	72.22%	(877350, 9514390)	Yes
(2,1)	Water	86.11%	(877410, 9514390)	Yes

*In the case of more than one modal class, the algorithm choses the smallest number. In order, Water refers to 1 and Forest refers to 2. So Water is selected.



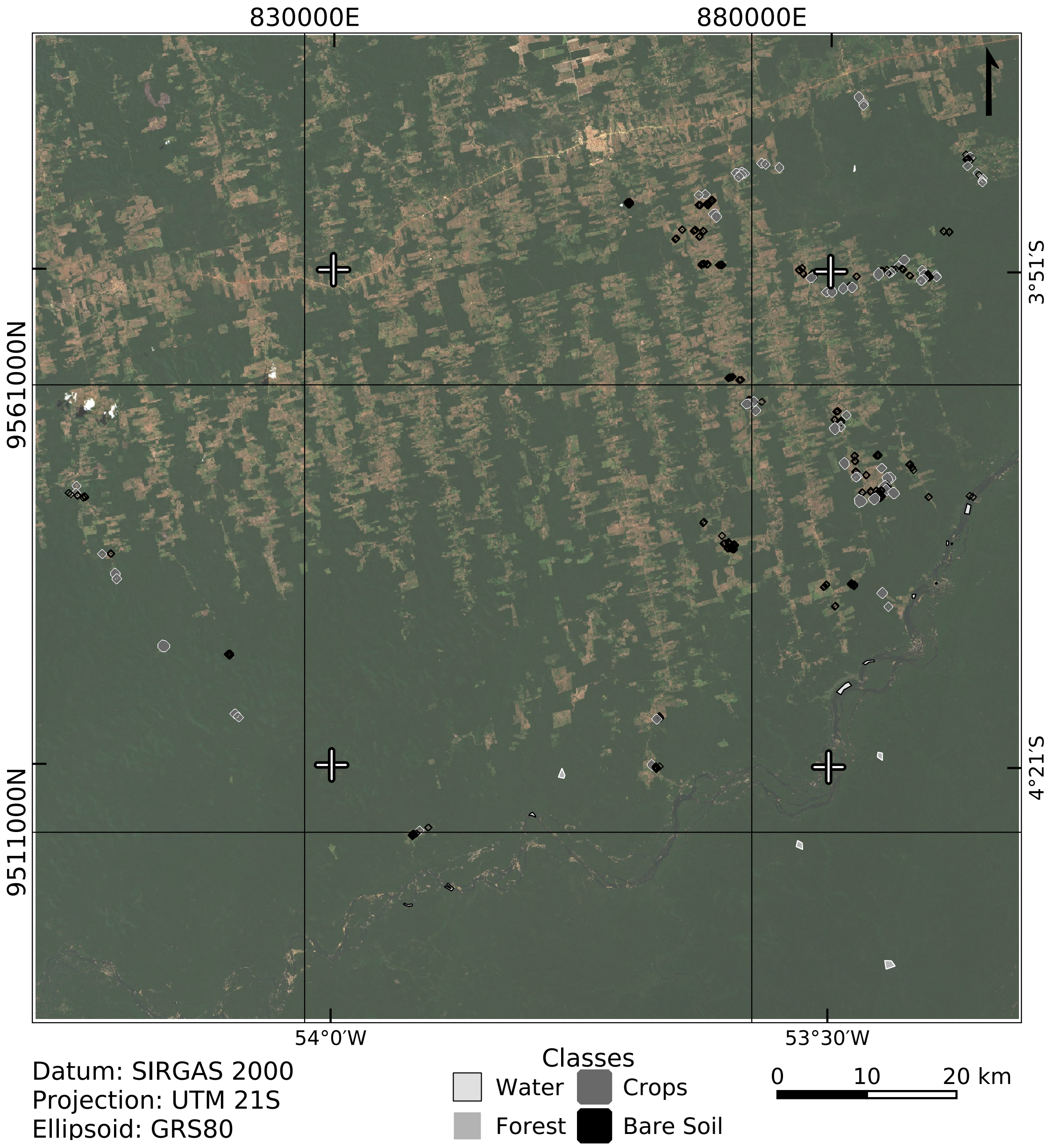
Cartographic Information

Datum: SIRGAS2000
Ellipsoid: GRS80
Coordinate Reference System: UTM 22S
Scale: Indicated

The grid mask runs through all LR pixels in order to acquire the modal class and the modal class proportion and add them to a geodataframe. (a) is the Thematic map or baseline image containing pixel row/column number and (b) is the attribute table of the geodataframe indicating which pixels are selected as candidate samples. This image in colours is Figure 4.3, referred to in Section 4.2.2, page 58.

SOURCE: Author.

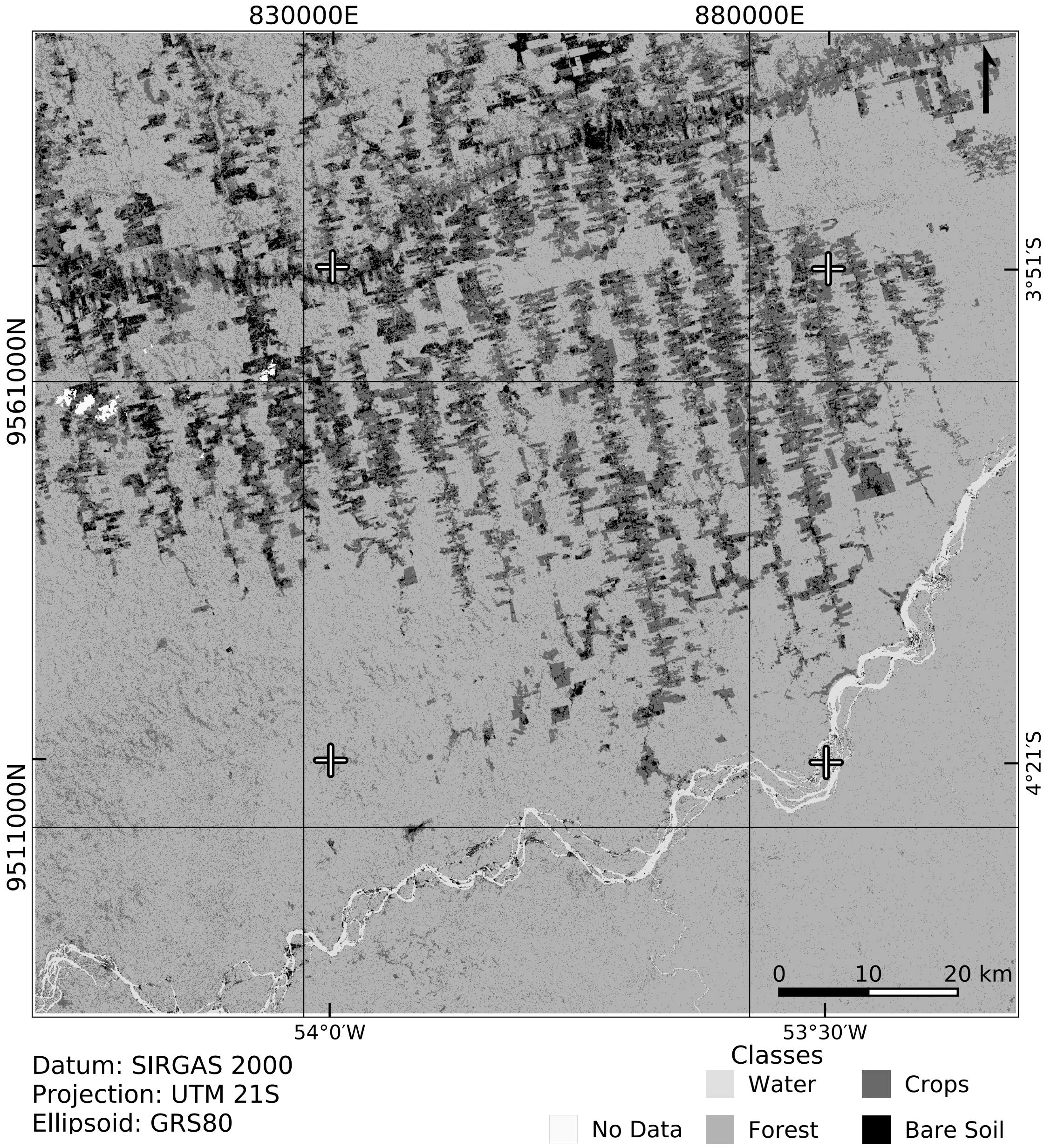
Figure C.2 - Pixel-based Baseline Classification using Random Forest - Region of Interest (ROI).



The polygons have been augmented for visualisation purposes. This image in colours is Figure 5.1, referred to in Section 5.1.1, page 66.

SOURCE: Author.

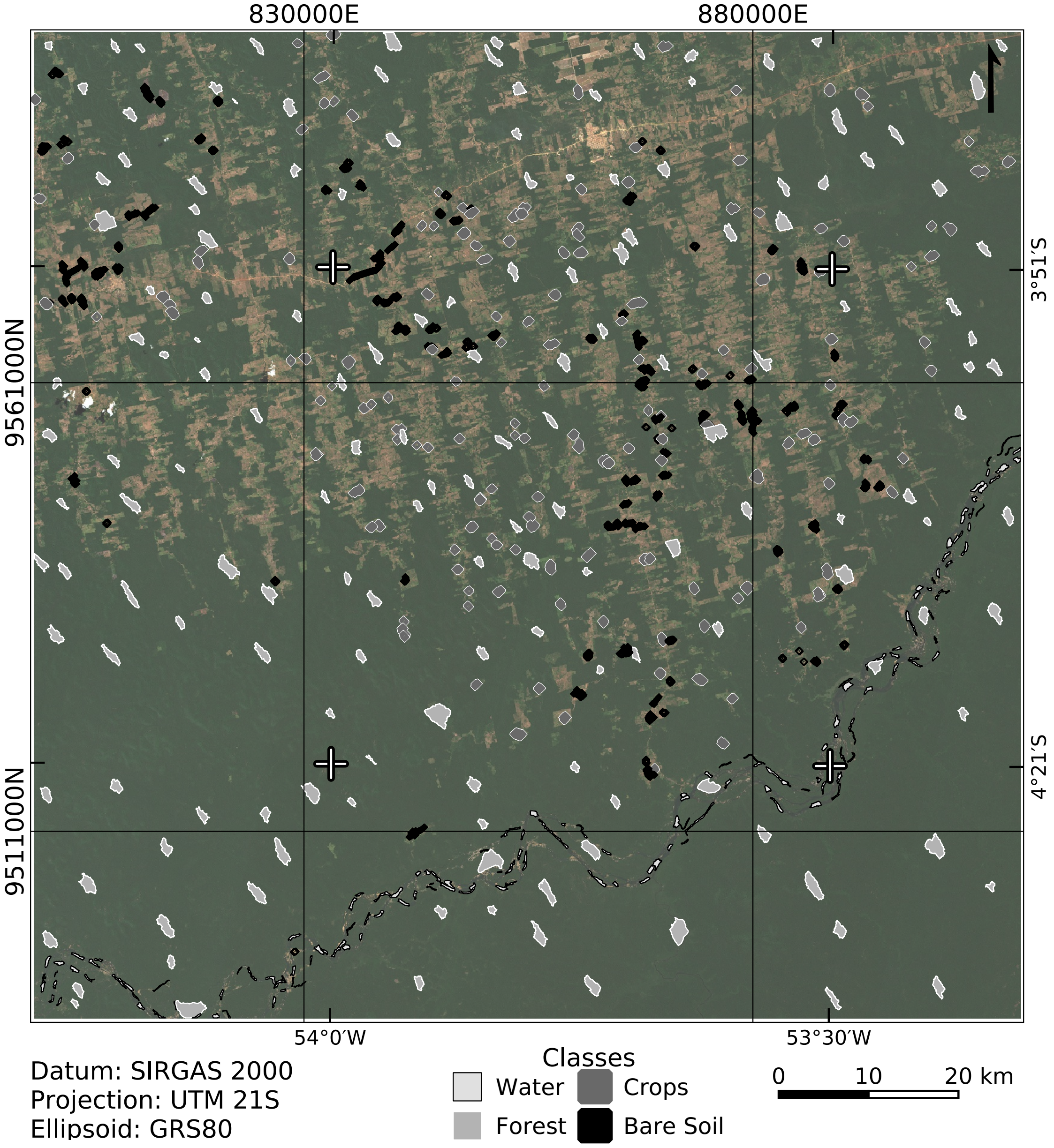
Figure C.3 - Pixel-based Baseline Classification using Random Forest - Classified Image in Black and White.



This image in colours is referred to in Section 5.1.1, page 66.

SOURCE: Author.

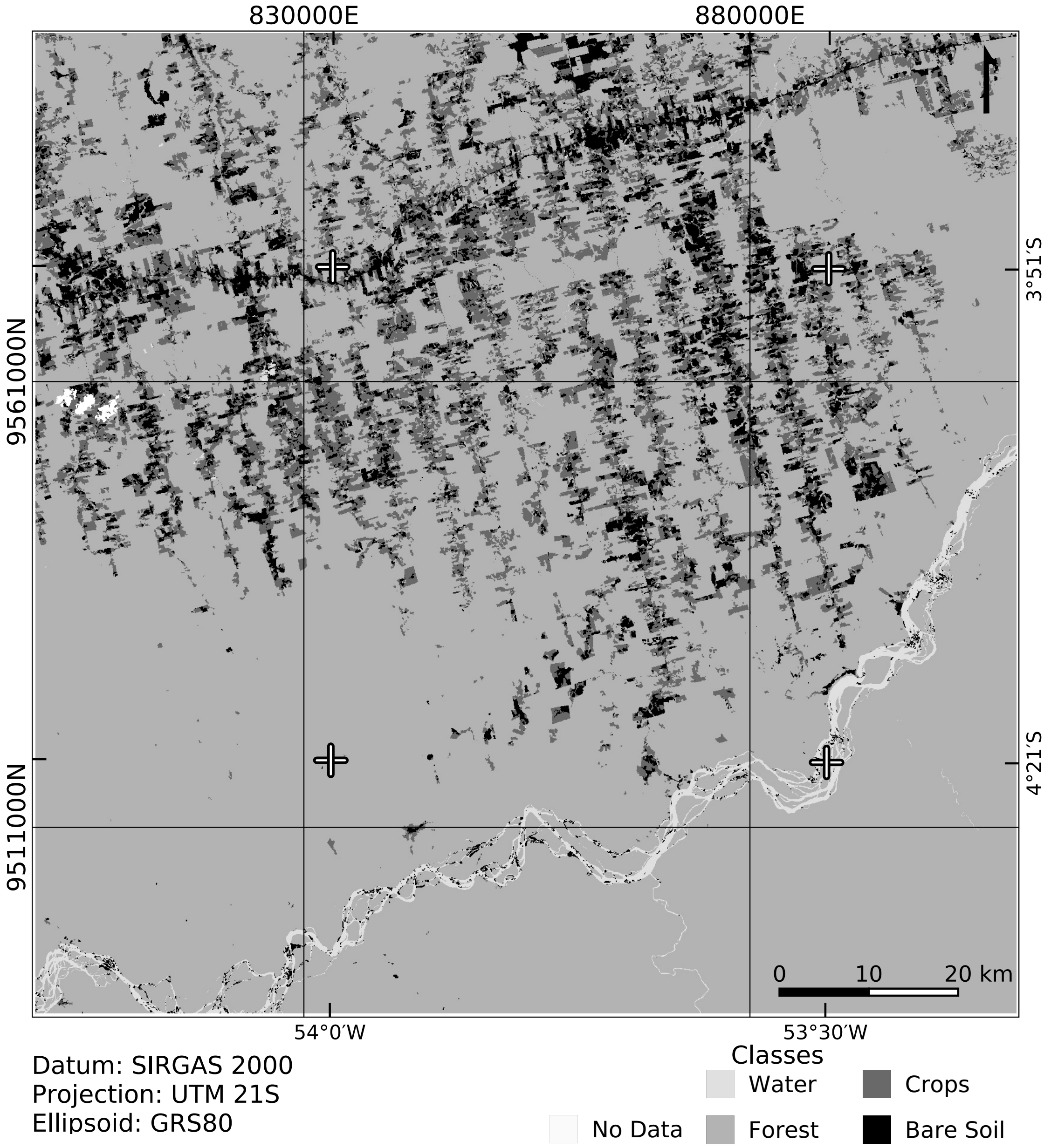
Figure C.4 - Region-growing Baseline Classification using Random Forest - Region of Interest (ROI).



The polygons have been augmented for visualisation purposes. This image in colours is Figure 5.8, page 88.

SOURCE: Author.

Figure C.5 - Region-growing First Classification using Decision - Classified Image in Black and White.



This image in colours is referred to page 90.

SOURCE: Author.

PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE

Teses e Dissertações (TDI)

Teses e Dissertações apresentadas nos Cursos de Pós-Graduação do INPE.

Notas Técnico-Científicas (NTC)

Incluem resultados preliminares de pesquisa, descrição de equipamentos, descrição e ou documentação de programas de computador, descrição de sistemas e experimentos, apresentação de testes, dados, atlas, e documentação de projetos de engenharia.

Propostas e Relatórios de Projetos (PRP)

São propostas de projetos técnico-científicos e relatórios de acompanhamento de projetos, atividades e convênios.

Publicações Seriadas

São os seriados técnico-científicos: boletins, periódicos, anuários e anais de eventos (simpósios e congressos). Constam destas publicações o Internacional Standard Serial Number (ISSN), que é um código único e definitivo para identificação de títulos de seriados.

Pré-publicações (PRE)

Todos os artigos publicados em periódicos, anais e como capítulos de livros.

Manuais Técnicos (MAN)

São publicações de caráter técnico que incluem normas, procedimentos, instruções e orientações.

Relatórios de Pesquisa (RPQ)

Reportam resultados ou progressos de pesquisas tanto de natureza técnica quanto científica, cujo nível seja compatível com o de uma publicação em periódico nacional ou internacional.

Publicações Didáticas (PUD)

Incluem apostilas, notas de aula e manuais didáticos.

Programas de Computador (PDC)

São a seqüência de instruções ou códigos, expressos em uma linguagem de programação compilada ou interpretada, a ser executada por um computador para alcançar um determinado objetivo. Aceitam-se tanto programas fonte quanto os executáveis.