Normalization of Remote Sensing Imagery for Automatic Information Extraction

A Master's Thesis
Submitted to the Faculty of
Escola Tècnica d'Enginyeria de Telecomunicació de Barcelona
Universitat Politècnica de Catalunya
by
Pol del Aguila Pla

In partial fulfilment
of the requirements for the degree of
TELECOMMUNICATIONS' ENGINEERING

Advisor: Felipe Calderero (UPF)
Second Advisor: Markus Flierl (KTH)

Barcelona, June 2014
Normalization of Remote Sensing Imagery for Automatic Information Extraction

Pol del Aguila Pla

March 17, 2014
For the time being, Remote Sensing automatized techniques are conventionally designed to be used exclusively on data captured by a particular sensor system. This convention was only adopted after evidence suggested that, in the field, algorithms that yield great results on data from one specific satellite or sensor, tend to underachieve on data from similar sensors. With this effect in mind, we will refer to remote sensing imagery as heterogeneous.

There have been attempts to compensate every effect on the data and obtain the underlying physical property that carries the information, the ground reflectance. Because of their improvement of the informative value of each image, some of them have even been standardized as common preprocessing methods. However, these techniques generally require further knowledge on certain atmospheric properties at the time the data was captured. This information is generally not available and has to be estimated or guessed by experts, a very time consuming, inaccurate and expensive task. Moreover, even if the results do improve in each of the treated images, a significant decrease of their heterogeneity is not achieved.

There have been more automatized proposals to treat the data in the literature, which have been broadly named RRN (Relative Radiometric Normalization) algorithms. These consider the problem of heterogeneity itself and use properties strictly related to the statistics of remote sensing imagery to solve it.

In this master thesis, an automatic algorithm to reduce heterogeneity in generic imagery is designed, characterized and evaluated through crossed classification results on remote sensing imagery.
This Master’s Thesis would not have been possible without the effort, organization and collaboration of a lot of people from four different universities. From ETSETB, UPC, professor Ferran Marqués put together and coordinated a team that stretched from Gran Canaria to Stockholm so this project could be done. From UPF, also in Barcelona, visiting professor Felipe Calderero supervised all of my work and helped me into the field of remote sensing. From ULPGC, in Gran Canaria, Canary Islands, professors J. Marcello and F. Eugenio, as well as their teams, provided me with the much needed datasets, and with careful and convenient advice when I needed it. From the school of Electrical Engineering at KTH, associate professor Markus Flierl, as well as some of the PhD students in the Communication Theory department, guided me through every administrative step and provided helpful tips and ideas.

Thank you, Marie Maros, for your support, both in personal and academic matters. Thank you for your time, your understanding, your patience and your help.

To my parents, Agustí and Georgina, and sister, Aina, thank you for believing in me and investing in my future.
# CONTENTS

## INTRODUCTION AND PRELIMINARIES

1. **INTRODUCTION**  
   1.1 Motivation  
   1.2 State of the art  
      1.2.1 LU/LC Classification  
      1.2.2 RRN algorithms  
   1.3 Goals  
   1.4 Thesis outline

2. **PRELIMINARIES**  
   2.1 Heterogeneity and how to evaluate homogenization  
   2.2 Remote Sensing  
   2.3 LU/LC classification methods and standards

## HOMOGENIZATION

3. **VARIANCE COMPREHENSIVE LINEAR REGRESSION TO THE MEAN**  
   3.1 Preconditions  
      3.1.1 Linear sensor model  
   3.2 Motivation by simple example  
   3.3 Extension and implementation issues  
      3.3.1 Dimensionality, Variance vs Frobenius norm of the Covariance Matrix  
      3.3.2 Offset: Forcing zero mean  
      3.3.3 Projecting to the mean and iterative behavior  
      3.3.4 Further extensions

4. **NON-LINEAR OUTLIER DETECTION AND REMOVAL**  
   4.1 Theoretical background  
      4.1.1 Rician modeling  
      4.1.2 Variances for outlier detection  
   4.2 Proposal

## CLASSIFICATION AND SEGMENTATION

5. **CLASSIFICATION: CHOICE OF CLASSIFICATION ALGORITHM**  
   5.1 Reduced dataset selection  
   5.2 Classifier selection  
      5.2.1 Parametric classifiers, parameter selection  
      5.2.2 Features  
      5.2.3 Results

6. **SEGMENTATION**  
   6.1 Texture information

---

**Page 4**
<table>
<thead>
<tr>
<th>Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.2 Active Contours, Chan &amp; Vese algorithm for segmentation</td>
<td>46</td>
</tr>
<tr>
<td>6.3 Combining results</td>
<td>47</td>
</tr>
<tr>
<td>6.3.1 Combining multiple segmentations</td>
<td>47</td>
</tr>
<tr>
<td>6.3.2 Combining segmentation and classification</td>
<td>48</td>
</tr>
<tr>
<td>4 EXPERIMENTS AND CONCLUSIONS</td>
<td>49</td>
</tr>
<tr>
<td>7 EXPERIMENTAL EVALUATION</td>
<td>50</td>
</tr>
<tr>
<td>7.1 Homogenization - RRN</td>
<td>50</td>
</tr>
<tr>
<td>7.1.1 Preprocessing, spatial resolution</td>
<td>50</td>
</tr>
<tr>
<td>7.1.2 6S radiometric correction as homogenization</td>
<td>55</td>
</tr>
<tr>
<td>7.1.3 RRN on Full Gran Canaria dataset</td>
<td>58</td>
</tr>
<tr>
<td>7.1.4 RRN on SPOT imagery and RRN on LANDSAT imagery</td>
<td>63</td>
</tr>
<tr>
<td>7.2 Segmentation</td>
<td>66</td>
</tr>
<tr>
<td>8 CONCLUSIONS</td>
<td>70</td>
</tr>
<tr>
<td>5 APPENDIX</td>
<td>74</td>
</tr>
<tr>
<td>6 RESULTS</td>
<td>75</td>
</tr>
</tbody>
</table>
Figure 1  Search on an exponential grid \((2^n)\) of the optimal parameters for the RBF kernel \((C, \gamma)\) on image C before any preprocessing. The choice of \(\gamma\) was found irrelevant, and the optimal \(C\) is 1024.

Figure 2  Images involved in the proportionality constant choice. As it can be seen in their details in Table 4, these images were captured with a difference of only 9 days. Note that, while the LU/LC information practically does not change, the diversity within them is high, with clouds, atmospheric issues, and different sensors.

Figure 3  Experimental choice of the proportionality constant between the scale change and the Gaussian kernel standard deviation. The chosen is finally \(k = 0.5\) according to crossed LU/LC classification results. The false color representation of the images involved can be seen in Figure 2, and their details found in Table 4.

Figure 4  Crossed LU/LC classification results on reduced dataset. Only images that were available both before and after the 6S radiometric correction [17, 16] were used, their details can be found in Table 5. A shows the results on the original images without preprocessing. B shows the results after the images were processed by Algorithms 3 and 4. C shows the results on the 6S corrected versions of the images. D shows the results after applying the developed algorithms to the 6S corrected versions of the images.

Figure 5  Crossed LU/LC classification results on the whole Gran Canaria dataset. Top, before RRN corrections. Bottom, after RRN by Algorithms 3 and 4. The details of the images involved can be seen in Table 6.

Figure 6  False color representation of the images with least satisfactory results in the crossed LU/LC classification test. The crossed classification results can be found in 5.
Figure 7  LU/LC classification of Image 16 when the classifier is trained on Image 9, before and after RRN by Algorithms 3 and 4. See Table 6 for details on the images.

Figure 8  Crossed LU/LC classification results in the SPOT images from the Gran Canaria dataset. Top, the results when Algorithms 3 and 4 are fed with all the images in the Gran Canaria dataset, from all satellites. Bottom, when the RRN is done using exclusively information from SPOT images. The image indices are the same that were used in Section 7.1.3. The image details can be found in Table 6.

Figure 9  Crossed LU/LC classification results in the LANDSAT images from the Gran Canaria dataset. Top, the results when Algorithms 3 and 4 are fed with all the images in the Gran Canaria dataset, from all satellites. Bottom, when the RRN is done using exclusively information from LANDSAT images. The image indices are the same that were used in Section 7.1.3. The image details can be found in Table 6.

Figure 10  Top, Elevation model used to find ground truth data for segmentation, bottom.

Figure 11  Segmentation error in the 21 images from the Gran Canaria dataset. The image details can be found in Table 6. Top, percentage of land pixels segmented as sea. Middle, percentage of sea pixels segmented as land. Bottom, mean error percentage.

Figure 12  Top, segmentation results superimposed in a false color representation of Gran Canaria. Bottom, classification of Image 16 when training on Image 10, after the RRN procedure. The images’ details can be found in Table 6. It is clear that in this case, the segmentation results could aid LU/LC classification.
LIST OF TABLES

Table 1  Spectral bands for different satellites. B, G and R stand for Blue, Green and Red. NIR, SWIR and LWIR correspond to the Near, Short Wave and Long Wave Infrared Regions. HSWIR is not standard notation and refers to the upper part of the SWIR band. The measure between brackets is the one corresponding to the side pixel size.  

Table 2  Description of the dataset used to obtain experimental background for the selection of a classifier. All images were radiometrically corrected by the 6S radiative transfer code before experimentation. Refer to Table 1 for specific information on each satellite’s properties.  

Table 3  Accuracy over the different databases for each combination image - classifier - feature. In all tests the classifiers were trained and evaluated on the spectrum and the feature after the + symbol, when applicable. Results on the training database are included for completion.  

Table 4  Details of the images involved in the proportionality constant choice. Their false color representations can be seen in Figure 2.  

Table 5  Details of the images available in both original and 6S corrected version.  

Table 6  Details of the images in the full Gran Canaria dataset.  

Table 7  Quantitative evaluation shown as a colored matrix in the bottom part of Figure 9. Corresponds to the crossed LU/LC classification accuracies for the LANDSAT images available. Data for the upper part of Figure 9 corresponds to the appropriate subset of Table 10.  

Table 8  Percent crossed LU/LC classification accuracies shown as colored matrices in Figure 4. Experiments A, B, C and D are defined in the original Figure 4. Further details may be read in Section 7.1.2.
List of Tables

<table>
<thead>
<tr>
<th>Table 9</th>
<th>Data represented in the upper part of Figure 5. Mean self classification accuracy: 84.4 %. Mean crossed classification accuracy: 42.4 %.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 10</td>
<td>Data represented in the lower part of Figure 5. Mean self classification accuracy: 84.3 %. Mean crossed classification accuracy: 62.8 %</td>
</tr>
<tr>
<td>Table 11</td>
<td>Quantitative evaluation shown as a colored matrix in the bottom part of Figure 8. Corresponds to the crossed LU/LC classification accuracies for the SPOT images available. Data for the upper part of Figure 8 corresponds to the appropriate subset of Table 10.</td>
</tr>
</tbody>
</table>

**LIST OF ALGORITHMS**

<table>
<thead>
<tr>
<th></th>
<th>Algorithm</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Example: Linear regression to the mean</td>
<td>28</td>
</tr>
<tr>
<td>2</td>
<td>Example: Variance comprehensive linear regression to the mean</td>
<td>29</td>
</tr>
<tr>
<td>3</td>
<td>Iterative multi-band variance comprehensive linear regression to the mean</td>
<td>34</td>
</tr>
<tr>
<td>4</td>
<td>Model-based non-linear outlier detection and removal</td>
<td>38</td>
</tr>
</tbody>
</table>
ACRONYMS

UNIVERSITIES

**ETSETB**  Escola Tècnica Superior d’Enginyeria de Telecomunicació de Barcelona (Telecommunication Engineering School of Barcelona)

**UPC**  Universitat Politècnica de Catalunya (Polytechnic University of Catalonia)

**UPF**  Universitat Pompeu Fabra (Pompeu Fabra University)

**ULPGC**  Universidad de Las Palmas de Gran Canaria (University of Las Palmas de Gran Canaria)

**KTH**  Kungliga Tekniska Högskolan (Royal Institute of Technology)

REMOTE SENSING TERMS

**LU/LC**  Land Use and Land Cover

**RN**  Radiometric Normalization

**RRN**  Relative Radiometric Normalization

**LANDSAT**  Land Satellites. Remote sensing project that has launched a whole family of satellites, named from LANDSAT 1 to 8. The LANDSAT project is a joint initiative between the United Stats (of America) Geological Survey and NASA.

**SPOT**  Satellite Pour l’Observation de la Terre (Satellite for the observation of Earth). Remote sensing project that has launched a whole family of satellites, named from SPOT 1 to 7. The SPOT project is a joint initiative between the french CNES (Centre national d’études spatiales, the belgian SSTC (Services fédéraux des affaires Scientifiques, Techniques et Culturelles) and the Swedish SNSB (Swedish National Space Board, Rymdstyrelsen).

**PIF**  PseudoInvariant Features, terminology used in $^{28}[27]$ to refer to points that do not vary their reflectance values over time

**NIR**  Near Infrared
List of Algorithms

**NDWI**  Normalized Difference Water Index

**MNDWI**  Modification of the Normalized Difference Water Index

**NDVI**  Normalized Difference Vegetation Index

**SIGNAL PROCESSING TERMS**

**NN**  Neural Networks

**SVM**  Support Vector Machines

**MSE**  Mean Squared Error

**PDF**  Probability Density Function
Part I

INTRODUCTION AND PRELIMINARIES
INTRODUCTION

This thesis targets readers with a MsC level on statistical signal processing. Specialists in image processing that have not explored the field of remote sensing may find that part of the decisions taken during the thesis do not follow the usual standards. Some clue procedural differences between generic image classification and remote sensing image classification can certainly disorientate when they are first encountered. Most of these are due to one of the following points,

1. The large amount of pixels in a remote sensing image
2. The high price of each of the images
3. The statistical heterogeneity between different remote sensing images.

For further details on any of these points, or clear statements on what they imply, refer to Chapter 2.

The following presents the topics that will be covered in this Master’s Thesis. In Section 1.1, the motivation behind the research work performed is clarified. Section 1.2 summarizes how this thesis stands within the field of remote sensing imagery radiometric normalization, or, how it will be referred inside this thesis, homogenization. Finally, Sections 1.3 and 1.4 ease the comprehension of this thesis as a whole by clearly specifying its goals and detailing the content of each of the following chapters, respectively.

1.1 MOTIVATION

LU/LC maps are geolocalized images that establish, for an specific time instant, the use given to each of the represented positions, or what was covering that point at that moment. Typical labels or classes in these maps are Wild Vegetation, Water, Bare soil, Agriculture or Urban area. Extracting this information from remote sensing imagery in an automated manner is a task that is not easily solved. As a matter of fact, even picking images from which is reasonable to attempt to extract LU/LC maps is a hard endeavor, as clouds and other disturbances can easily invalidate large amounts of data.

Developing methods that automatically extract this information from images taken at different dates is even more difficult, because,
even if only one sensor is used, its intrinsic properties may have changed due to degradation in time \cite{28}. In fact, the problem of having two images from the same sensor at different times, or two images from two sensors with the same spacial and spectral resolution, are generally indistinguishable and treated with the same procedures \cite{21}.

As a consequence to all these random differences, satellites are generally designed with specific purposes in mind, and specifically tweaked to ease processing of a certain kind. Ironically, this fact makes it even more difficult to develop algorithms capable of automatically generating general LU/LC maps at different time points to study the evolution of land’s use and cover, this is, if one wants to be able to use information from any satellite. Most, if not all, modern studies performed on remote sensing imagery rely only on data from a particular sensor, or sensor family \cite{3,19}.

Furthermore, as the design of a satellite generally entails a lot of complicated and budget limited decisions, even satellite families designed with the same purpose have very different properties. Table 1 in Chapter 2 contains the specifications of some satellites from two of the most used satellite families, both intended for the general observation of Earth. There one can easily see that even bands with the same denomination, or that are supposed to cover the same band, do not have the same technical definition.

The motivation behind this thesis is to provide a way to bring all remote sensing imagery in a temporal study of the same area to a common space, where statistical analysis algorithms are guaranteed to be run with equivalent results. Of course, this aim is extremely broad and ambitious, and has partly already been addressed by the previous literature. Section 1.2 summarizes what has been published on this matter up to this date, and Section 1.3 constraints what will be attempted within this Master’s Thesis’ project.

1.2 STATE OF THE ART

This thesis will cover material coming from two different, if not disjoint, research fields within remote sensing. As it is, most of the work will be related to the field of Relative Radiometric Normalization (RRN), a set of techniques or algorithms that attempt to ease comparison and evolution studies on remote sensing imagery. However, the targeted application and the ultimate test-bench of the thesis is Land Use and Cover (LU/LC) classification, that has been selected by its undeniable intrinsic value and the implications of its success. The following summarizes which is the current state of the art in both these fields and establishes in which points this thesis tries to diverge from it.
1.2 STATE OF THE ART

1.2.1 LU/LC Classification

As in any other classification problem, there are broadly two different choices within the design of a system to classify remote sensing images and generate LU/LC maps. First of all, the input feature space has to be chosen to maximize the separability of the targeted classes. Afterwards, a classification method capable of defining borders flexible enough to separate the resulting clusters must be found.

In LU/LC classification, a lot of work has been done to find the best feature space. Historically, non-linear projections or normalized indices have been defined for each class \cite{26,23,35}. These are not derived statistically or automatically for each dataset, but based on theoretical properties of the physical surface and radiation theory. This fact makes them theoretically consistent, extremely informative, and allows interpretations of their values \cite{13}. In fact, studies such as \cite{9} proof the possibility of extracting relevant evolution information directly form indices, instead of generating LU/LC maps and then interpreting them. However, none of these changes the fact that the indices are not designed to provide the best separability between classes, but to exploit specific physical properties. Despite this fact, current LU/LC classification studies still consider them as a basis for classification, frequently arriving to opposite conclusions \cite{32,19}.

Other features that are frequently mentioned are those based on textural information, such as the local entropy within a window, the local variance within a window or even multi-scale information \cite{3,20}.

When it comes to the selection of a classification algorithm, those based on neural networks (NN) or support vector machines (SVM) are generally better at the task at hand, independently of the chosen features. The non-linearity that can be introduced in both families of classifiers allows them to learn the irregular borders that separate the different LU/LC classes. By studying the current literature, in fact, it appears that all methods related to the kernel trick yield better results in the field of remote sensing \cite{24,22}.

In the following thesis, accordingly with the observations in \cite{15}, SVM classifiers with Radial Basis Functions (RBF) kernels will be used for every LU/LC classification. The RBF kernel parameters will be optimized through the experimental procedure also proposed in \cite{15}. Additionally, segmentation techniques from \cite{6} will be used to aid classification and exploit spatial and textural information.

1.2.2 RRN algorithms

The first satellite of the LANDSAT family was launched in 1972. Ever since, Radiometric Normalization has been an open research field. The first proposals appeared in the late 1980s and early 1990s. Reference \cite{34} provides a comprehensive comparison of these early tech-
Techniques, that are mainly divided between model-based and experimental.

Model-based techniques require ancillary information that has to be generated manually or semi-automatically by experts, and generally target the recovery of the actual radiance instead of the projection to a new common space. Only one model-based technique has remained and is now part of the common preprocessing steps, that is, the 6S radiative transfer code for atmospheric correction of satellite data, which was consistently validated in [17][16].

The experimental, data-driven or Relative algorithms mostly represent different expressions of the same intuition. All of them agree that the difference between the sensors’ response is linear, and propose different linear regressions to correct it. The parameters of this linear regression are always estimated using areas that are considered to have the same reflectance in all the obtained samples. It is only on the assigned terminology for this locations and the methodology to select them where there are most differences among the proposals. References [28] and [27] refer to these positions as pseudo-invariant features (PIF), but basically search for urban related spectral responses, as they consider them to be less prone to change in time. [12] uses two sets of points, named dark and bright control sets, which are found through the use of scatterplots of the Kauth and Thomas greenness and brightness projections [14]. [33],[11] locate their points, which they refer to as no-change pixels (NC), by using the scatterplots between the near-infrared bands of each image to locate stable water and land pixels. Once these locations are obtained, all proposals derive the linear regression parameters, $(\alpha_{n,j}, \phi_{n,j})$, for each band $j = 1, \ldots, N_{bands}$ and image $n = 2, \ldots, N_{images}$, as a function of the mean, variance, minimum or maximum value the specific image and band had over those positions. Note that these regressions are usually intended to bring all images to one. This last image is generally referred to as control image and corresponds to $n = 1$ above.

The field of RRN has evolved since the publication of these first articles, and the use of the terms no-change pixels and pseudoinvariant features has prevailed. [10],[4],[1] are examples on how researchers have attempted to make the selection of these special locations less subjective. However, [25], an extremely recent study on RRN techniques for change detection studies, still proposes a semi-automatic methodology to solve the task at hand. Moreover, because of the same property of remote sensing imagery these techniques are trying to deal with, heterogeneity, it is not uncommon to see publications obtaining opposite conclusions on their respective datasets. The appropriate measure to quantify each algorithm’s performance has not been properly defined, and the number of images each study is based is generally insufficient.
In the early studies, remote sensing imagery was both expensive and hard to obtain. However, at the current time, many images can be obtained of practically any region in the world, provided for free by the LANDSAT project. \cite{28} studies two different scenarios, for which three and two images are used respectively. In \cite{12} a total of six images is used, while \cite{34} compares six different RRN algorithms by using a total of two images. \cite{30} uses seven images, and \cite{10} only three. \cite{4} studies again two different scenarios, for which five and two images are used respectively. The most recent study, \cite{25}, uses only five images. A common measure of performance, used in \cite{28,12,34,1} is the mean squared error (MSE) between the values of the PIFs after the RRN and their values in the control image. However, the PIFs or no-change points are defined subjectively within the same algorithm that is being evaluated. Moreover, the MSE measure is not consistent with any of the approaches, because none of them propose to obtain the linear regression parameters by optimizing the MSE over the no-change locations. \cite{30} takes an arguably more specific but consistent approach to performance assessment and uses crossed classification results, both from control image to corrected image and vice-versa.

In this thesis, an automatic RRN algorithm will be proposed and thoroughly described. The selected performance measure will be crossed classification from corrected image to corrected image, as the proposal will not rely on a control image, but will project all images to a new, middle-ground space. In reference \cite{34}, it was mentioned that RRN algorithms that use bigger no-change sets perform better in the overall image. Following this lead, the introduced proposal will use every pixel in the image, combined with an estimation of its variance in time, to find the best linear regression. The experimentation will be performed in 21 different images in the region of Gran Canaria, in the Canary islands, near to the north-western African coast.

1.3 Goals

During this thesis it will be mentioned repetitively that remote sensing imagery is heterogeneous. The precise mathematical definition that will be used for this is stated in Chapter 2, but it can be understood intuitively. In remote sensing images, the same kinds of land cover or use are not characterized the same way in different images. There is, of course, a global tendency that allows to pinpoint most phenomena with the naked eye. However, the changes in color, shade, clarity and contrast make it harder to decide without a broader context, and make some pixel to pixel decisions extremely hard.

Different remote sensing images vary in the following properties,

- Pixel size or spatial resolution,
1.4 Thesis Outline

- Number of spectral bands,
- Definition of the spectral bands with the same names,
- Sensor responses,
- Capture conditions, including but not limited to,
  - Atmospherical conditions,
  - Illumination distribution,
  - Weather conditions.

These thesis’ goals are all aimed to reduce, correct, or compensate the heterogeneity in groups of remote sensing images to allow their comparison and use in evolution or change detection algorithms. However, it would be unreasonable for a Master’s level thesis to target to solve all the heterogeneity above. With respect to this, the pixel sizes will be assumed the same, and when they are not, the Gaussian scale space will be used to force it. The number of spectral bands will be considered the same, and only those bands present in all the considered satellites will be used. Finally, the definition of the spectral bands, following the lead of the stat of the art, will be considered identical.

It has to be mentioned that some of the following goals are also related to enhancing LU/LC classification results, the basis on which most change detection studies are carried out.

In short, this thesis’ goals are,

- To design an algorithm to homogenize several remote sensing images of the same region, with the same spatial resolution and bands,
- To explain how this algorithm fits within the field of RRN,
- To reason why crossed classification is a valid measure to evaluate homogenization,
- To evaluate this algorithm using crossed classification results,
- To provide a strategy to eliminate common disturbances in change detection studies, such as clouds and water reflections,
- To propose an algorithm to exploit textural and spatial information in change detection applications based on LU/LC classification in areas with large extensions of water.

1.4 Thesis Outline

This Chapter has introduced the motivation, contextualization and goals of this thesis and explains how it is structured.
Chapter 2 provides the background knowledge required to understand the design decisions taken in the developed algorithms. Specifically, Section 2.2 details the particularities of remote sensing imagery, and Section 2.3 exposes their consequences on usual classification procedures.

Chapters 3 and 4 present, motivate and clarify the main algorithm proposed in this thesis. Specifically, Chapter 3 presents a proposal for a RRN algorithm which will be referred as Variance Comprehensive Linear Regression to the mean, and Chapter 4 explains how it can be extended to detect clouds and other strong disturbances and soften their effect.

Chapter 6 presents a method to exploit textural and spatial information in order to improve LU/LC classification results through a more consistent detection of land and sea. To conclude the theoretical load of the thesis, Chapter 5 presents an experimental procedure to select the best SVM RBF classifier for the experimental section, and exposes its results on a reduced dataset.

Finally, Chapters 7 and 8 present the experimental results and the conclusions that can be drawn from them, and evaluate whether each goal in 1.3 has been achieved.
Preliminaries

2.1 Heterogeneity and How to Evaluate Homogenization

Data driven algorithms, such as classifiers, rely on certain assumptions. The most basic of these, is that the information or features to be processed are characterized by the same PDFs that have been learned from the training information. When this is not the case at hand, this thesis will refer to the data as heterogeneous. Note that the definition of heterogeneous data is extremely broad, since two identical datasets with no difference but an offset would fulfill it.

At this point there are two reasons that motivate doing a slight modification on the definition of heterogeneity. First of all, its definition should be related to intrinsic properties, and not be broken by just uninformative changes like the one described above. Second, the definition of heterogeneity should provide a way to measure it, because after all, the main goal will be to reduce it.

Suppose a dataset $A$ in which there are $N$ different classes of points. This is, if $A_k$ is the set of points of class $k$ inside $A$, $A = \bigcup_{k=1}^{N} A_k$. Each class also has a PDF $f_k(x_A): A \rightarrow \mathbb{R}$. Then, assume that this dataset is normalized, as it is usually done before classification, by subtracting the mean and dividing by the variance in each dimension. Over the new dataset, $\tilde{A}$, the PDF of each class is now $\tilde{f}_k(x_{\tilde{A}}): \tilde{A} \rightarrow \mathbb{R}$. Suppose now that the goal is to measure the heterogeneity between $A$ and a new dataset $B$, that includes the same $N$ classes. It will be much more representative to compare $\tilde{f}_k(x_{\tilde{A}})$ and $\tilde{f}_k(x_{\tilde{B}})$, than to compare the actual $f_k(x)$s. The new definition of heterogeneous data will therefore be analogous to the previous one, but over the $\tilde{f}_k(x)$s instead of over the $f_k(x)$s.

However, measuring heterogeneity under this definition is not trivial. Even when the class labeling for a whole dataset is known, i.e. the sets $A_k$ are completely known, estimating the PDFs is not a solved nor easy task. Moreover, if we consider the specific case of Remote Sensing and LULC classes, having even a small part of each $A_k$ is considered difficult. Classification, here, offers an extraordinary opportunity. By deriving the borders between the classes, and optimizing the performance, certain information about the unknown distribution probabilities is obtained. In fact, we know that the optimal performance is obtained when the border between two classes $k$ and
2.2 REMOTE SENSING

The term Remote Sensing broadly refers to a set of technologies and techniques that allow the measurement of magnitudes and phenomena from a distance.

The term is however generally used to refer to only those that are dedicated to measure physical magnitudes or detect phenomena on Earth, from the higher layers of the atmosphere or from space.

It is in this second sense in which this thesis will be based on remote sensing data. In particular, the main data source will be a set of satellital images of the Gran Canaria island. This is, one of the islands on the Canary Islands archipelago, next to the north-western African coast.

As it was announced in Section 1.1, remote sensing imagery has diverse or heterogeneous properties. Table 1 exemplifies how diverse the nominal properties of the sensors can be, differing in everything from spatial and spectral resolution to the definition of the same bands. Moreover, when it comes to the actual sensors’ responses, the differences are even bigger. Of course, there is also a different level of thermal noise for every system and every instant.

From the environment point of view, the differences can even be more important. The illumination is clearly not homogeneous across one image and depends extremely on the angle of the sun. The atmospheric phenomena, such as haze, can affect the data greatly. Even worse, there are big disturbances, generated either by meteorologic conditions or water reflexes.

In short, Remote Sensing data is incredibly diverse, and the common assumption within the field is that two images do not generally share statistical properties, even when they come from the same satel- lute and close time instants.

2.3 LU/LC CLASSIFICATION METHODS AND STANDARDS

Consistently with the high level of variation specified above, and the extreme cost of generating ground truth databases, some standards in
both the operation and the evaluation of LU/LC classification differ significantly from those of generic image classification.

Specifically, in LU/LC applications, it is usual to always train and test your algorithms on the same image. This does not break the so well-known restriction derived from pattern recognition theory, i.e. that one must keep separate test and train databases. The classifiers are usually trained in only a few points, and then tested in only a few, different, points. This implies that numerical or quantitative information is generally always contrasted with qualitative, intuitive impressions. In this thesis, an effort to evaluate the proposed methods numerically has been done. A special effort was also made by the people generating the databases in ULPGC, and the whole TELECAN project, to provide a ground truth as accurate and useful as possible.

The level of difficulty behind the generation of this ground truth data makes it a common practice, also, to pick only points that do not change their LU/LC classification over time. In this thesis, for all 21 images included, only one training and test database are used. These databases refer to points in space that do not change their classification during the studied period, and label them with a class. When the database has to be used, the values a specific image has in those positions are read, and used as training database. Even when this might seem irregular or flawed, this is actually not the case. If the points are picked diverse enough, there is no reason to assume that the fact that their classification does not change over time means they

<table>
<thead>
<tr>
<th>Band</th>
<th>SPOT 1 &amp; SPOT 2</th>
<th>SPOT 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>0,50µm − 0,59µm (20 m)</td>
<td>0,50µm − 0,59µm (20 m)</td>
</tr>
<tr>
<td>R</td>
<td>0,61µm − 0,68µm (20 m)</td>
<td>0,61µm − 0,68µm (20 m)</td>
</tr>
<tr>
<td>NIR</td>
<td>0,78µm − 0,89µm (20 m)</td>
<td>0,79µm − 0,89µm (20 m)</td>
</tr>
<tr>
<td>SWIR</td>
<td>×</td>
<td>1,58µm − 1,75µm (20 m)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Band</th>
<th>SPOT 5</th>
<th>Landsat 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>×</td>
<td>0,45µm − 0,52µm (30 m)</td>
</tr>
<tr>
<td>G</td>
<td>0,50µm − 0,59µm (10 m)</td>
<td>0,52µm − 0,60µm (30 m)</td>
</tr>
<tr>
<td>R</td>
<td>0,61µm − 0,68µm (10 m)</td>
<td>0,63µm − 0,69µm (30 m)</td>
</tr>
<tr>
<td>NIR</td>
<td>0,79µm − 0,89µm (10 m)</td>
<td>0,76µm − 0,90µm (30 m)</td>
</tr>
<tr>
<td>SWIR</td>
<td>1,58µm − 1,75µm (20 m)</td>
<td>1,55µm − 1,75µm (30 m)</td>
</tr>
<tr>
<td>HSWIR</td>
<td>×</td>
<td>2,08µm − 2,35µm (30 m)</td>
</tr>
<tr>
<td>LWIR</td>
<td>×</td>
<td>10,4µm − 12,5µm (30 m)</td>
</tr>
</tbody>
</table>

Table 1.: Spectral bands for different satellites. B, G and R stand for Blue, Green and Red. NIR, SWIR and LWIR correspond to the Near, Short Wave and Long Wave Infrared Regions. HSWIR is not standard notation and refers to the upper part of the SWIR band. The measure between brackets is the one corresponding to the side pixel size.
can not represent the whole class. In short, the points are picked so their PDF is independent to the fact that they do not change their LU/LC classification over time.
Part II

HOMOGENIZATION
This Chapter presents and theoretically discusses this thesis’ proposal of an algorithm that linearly homogenizes remote sensing imagery.

Algorithm 3 constitutes a novel approach to RRN. Its design has been done assuming a set of preconditions, tightly related to the established goals (see Section 1.3), and clearly specified in Section 3.1. The proposed algorithm in its final form is rather complex, mainly because of the high dimensionality of the problem at hand. Section 3.2 motivates its conceptual validity from a simple example with reduced dimensionality. Some considerations on how the algorithm was extended to a higher dimensionality and implemented are derived in section 3.3. Additionally, the assumed model on the different satellites’ sensors is discussed in 3.1.1. Finally, the iterative behavior of the algorithm is motivated and discussed in section 3.3.3.

3.1 Preconditions

First of all, all images should have the same nominal $N_b$ spectral bands. In other words, if some image covers a wider spectral range than another, only that part covered by both will be used by Algorithm 3. As it can be seen observing Table 1, this precondition implies that, if we are to use our whole set of images, only the Green, Red and Near Infrared bands will be used. As it will be further explained in Section 3.3.4, Algorithm 3 could easily be extended to use images with different bands, given certain requisites. However, this falls outside of the scope of this thesis.

Secondly, all images must have the same resolution. As it was mentioned in Section 1.3, once the data has been homogenized, it will be possible to exploit the interrelation between the images to generate higher resolution versions of them all. To do the actual homogenization, however, lower resolution versions generated through the Gaussian scale space are used.

Finally, all images will need to be geolocalized or rectified, in a way that there is a one to one correspondence between their pixels’ locations. Specific methods to do so are not proposed, as remote
sensing images are commonly provided already according to these needs.

Even though there is no precondition on the number of images used, it will be seen in Section 3.2 that the algorithm uses the estimations of both first and second order moments of each pixel through time. Obviously, the quality of these estimates will depend on the number of samples, i.e. the number of images included in the specific change detection project in which the algorithm is used.

3.1.1 Linear sensor model

The sensors have been considered to follow a linear model, as supported by the literature in the field of relative radiometric normalization. [28] is one of the most referred articles to reason this linear assumption, but the consistency with which this assumption has been made during the evolution of the field is the most convincing evidence (see [12, 30, 10, 4, 21, 11, 25]).

Explicitly, this assumption means that if the same pixel was captured under the same circumstances by two different sensors \( s_1 \) and \( s_2 \) with the same spectral resolution and number of bands \( N_b \), obtaining two measurements \( y_1, y_2 \in \mathbb{R}^{N_b} \), these would be related by Equation (2), if random effects were disregarded. Note that the offset \( \phi \in \mathbb{R}^{N_b} \) also has a different value for each band.

\[
y_1 = \begin{bmatrix} \alpha_1 & 0 & \ldots & 0 \\ 0 & \alpha_2 & \ldots & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \ldots & \alpha_{N_b} \end{bmatrix} y_2 + \phi
\]  

3.2 Motivation by simple example

The basic idea behind Algorithm 5 can be derived from the following example.

Assume a set of gray-scaled images \( A_n \in \mathbb{R}^{N_y \times N_x} \) for \( n = [1, \ldots, N_i] \), where \( N_y \) specifies the number of rows, \( N_x \) the number of columns, and \( N_i \) the number of images. Assume also that each of them was captured by a sensor \( s_n \) with a sensibility to light \( \gamma_n \), an offset \( \beta_n \) and some noise characterization \( f_n \). If these images \( \{A_n\} \) all picture the same scene under different lighting and environmental conditions, it is at this point clear that a classifier trained to recognize certain regions or classes within one of the images would not necessary work on another one. The random variations included by the different illumination, sensor and environment will easily change the distribution of each class, effectively making the set of images \( \{A_n\} \) statistically heterogeneous.
Without any further information on the disturbances created by the illumination and environmental changes, a sensible approach would be to try to partly cancel the effect of the sensor differences. This would apparently require estimating the sensors’ parameters $\gamma_n$ and $\beta_n$. This is, in turn, very hard without any calibration information, a robust estimation of the noise parameters $\{f_n\}$, or ground truth knowledge on each of the actual light intensities.

Taking into account that the goal is to bring all the images to a common ground and not to recover the measured light intensities $Z_n = \frac{1}{\gamma_n} (A_n - \beta_n 1)$, one realizes that the estimation of $\{\gamma_n, \beta_n\}$ is not actually necessary. It would suffice to find a linearly dependant space where all images can be represented equally, and then find the appropriate projectors to bring each image there. In other words, a first bold proposal to compensate for the sensors’ differences would be Algorithm 1.

Algorithm 1 Example: Linear regression to the mean

1: $\bar{A} = \frac{1}{N_i} \sum_{n=1}^{N_i} A_n$
2: for $n = 1$ to $N_i$ do
3: $\{\hat{\alpha}_n, \hat{\phi}_n\} = \text{argmin} \left\{ \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} (\alpha_n A_n(x,y) + \phi_n - \bar{A}(x,y))^2 \right\}$
4: $A_n = \hat{\alpha}_n A_n + \hat{\phi}_n 1$
5: end for

Notice, however, that this would be equivalent to disregarding completely the fact that there are changes in the images produced by other than just the differences between the sensors $\{s_n\}$. Doing so would most probably result in ill-chosen projectors $\{\alpha_n, \phi_n\}$, more related to these other disturbances than to the sensors’ properties. In fact, both in this example and in this thesis’ application case, remote sensing images, illumination and environmental effects produce much higher variations between images than the sensor differences.

The main idea behind Algorithm 3 is to modify the cost function in line 3 by weighting each pixel’s square difference with the inverse of the pixel’s variance through the different images. Because the differences between the sensors will affect all pixels, the pixels that only vary because of them will be the ones with lower variance and higher weight. Therefore, by using these weights in the cost function, we establish that the effort done on minimizing the quadratic difference on line 3 has to be, for each pixel, proportional to our belief in the fact that it does not contain spurious effects. Note that, if we consider this proposal within the state of the art on RRN algorithms (see Section 1.2.2), it is only a natural soft extension of the NC or PIF approaches, i.e. an approach that avoids hard decisions between what is an invariant pixel and what is not.

28
Summarizing, Algorithm 3 is only an iterative, multi-band, refined version of the idea in Algorithm 2.

**Algorithm 2** Example: Variance comprehensive linear regression to the mean

\[ \tilde{A} = \sum_{n=1}^{N_i} \frac{1}{N_i} A_n \]

2: for all \((x, y) \in [1, N_x] \times [1, N_y]\) do

\[ \sigma^2(x, y) = \frac{1}{N_i-1} \sum_{n=1}^{N_i} (A_n(x, y) - \tilde{A}(x, y))^2 \]

4: end for

for \(n = 1\) to \(N_i\) do

6: \( \{ \hat{a}_n, \hat{\phi}_n \} = \arg\min \left\{ \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} \frac{1}{\sigma^2(x, y)} (a_n A_n(x, y) + \phi_n - \tilde{A}(x, y))^2 \right\} \)

8: end for

Note that this method holds for a wide variety of scenarios in which one wants to linearly bring different datasets to a common ground, and does not have a good modeling of the random effects they contain.

Furthermore, suppose that the previous example is extended, and now the images \(\{ A_n \}\) include some regions with changes that want to be kept and detected after the homogenization process. It is easy to see that following Algorithm 2 would avoid that these changes were considered part of the sensors’ variations, as long as they produced regions with higher \(\sigma^2(x, y)\). The results in Chapter 7 will provide results that suggest that this is so in our application problem. Finally, notice also that, due to its design, this algorithm is robust to clouds and other disturbances with similar behaviors, because of their strong effect on the variance \(\sigma^2(x, y)\).

3.3 Extension and Implementation Issues

There are obvious differences, however, between Algorithms 2 and 3. Even further, there are important differences between Algorithm 3 and the efficient implementation that has been used to generate all the results in Chapter 7.

In the following, the theoretical implications of most steps from Algorithm 2 to 3 are discussed, and some remarks on how Algorithm 3 was finally implemented are given.
3.3 Extension and Implementation Issues

3.3.1 Dimensionality, Variance vs Frobenius norm of the Covariance Matrix

As explained in section 3.2, the weight \( \frac{1}{\sigma^2(x,y)} \) in line 6 of Algorithm 2 intends to control the effect of each pixel in the optimization of the linear regression parameters \( \{\alpha_n, \phi_n\} \), according to its estimated variance across the images.

In real Remote Sensing data, however, each pixel is a vector with \( N_b \) bands, and each image requires the estimation of two vectors, \( \alpha_n \) and \( \phi_n \). This fact leads to a decision between the two most intuitive strategies to quantify the multi-band variation, both representing one extreme assumption or another on the level of band correlation of the unwanted variations. In the first place, one could assume that the disturbances present in the data are band-wise independent. That would mean that each pair of elements \( (\alpha_{n,j}, \phi_{n,j}) \) should be estimated using a band-wise variance \( \sigma_j^2(x,y) \), which would have been estimated using information from all the \( j \) bands. On the other hand, if the assumption is that the disturbances are highly band-wise correlated, one would like to define a measure of the amount of global variation in each pixel.

In this application case, the disturbances against which the estimation of \( \alpha_n \) and \( \phi_n \) has to be protected happen to be extremely correlated. These will mainly consist of clouds and reflection artifacts over large water areas, as well as extreme changes in the orography of the area. Therefore, a measure of global variation that happened to be convenient and is arguably theoretically appropriate was selected, the Frobenius norm of the covariance matrix of each pixel. Its convenience, in fact, allows to implement the algorithm independently of the number of bands, in a way that makes line 6 of Algorithm 2 just a particular case for \( N_b = 1 \). The new weights are referred in algorithm 3 and this section as \( \Omega(x,y) = \frac{1}{\Sigma(x,y)} \), where \( \Sigma \) is defined in Equation (3).

\[
\Sigma(x,y) = \| \hat{\mathbf{C}}_{x,y} \|_F = \sqrt{\text{tr} \left\{ \hat{\mathbf{C}}_{x,y} \hat{\mathbf{C}}_{x,y}^H \right\}} \quad (3)
\]

In terms of efficiency and implementation, it is at this point clear that computing the whole covariance matrix for each pixel and then extracting its Frobenius norm is not the most efficient way to obtain the weights \( \Sigma \) needed to estimate \( \alpha_n \) and \( \phi_n \). Taking into account the covariance matrix estimation \( \hat{\mathbf{C}}_{x,y} \) expression in the derivations in

\[1\] See Equation (3)
3.3 EXTENSION AND IMPLEMENTATION ISSUES

(3) show how this variation map can be computed in polynomial time with \( N_i \) and \( N_b \).

\[
\hat{C}_{x,y} = \frac{1}{N_i - 1} \sum_{n=1}^{N_i} (I_n(x,y) - \bar{I}(x,y)) (I_n(x,y) - \bar{I}(x,y))^H
\]

(4)

\[
\Sigma(x,y) = \frac{1}{N_i - 1} \sqrt{\begin{pmatrix} \text{tr} \left\{ \left( \sum_{n=1}^{N_i} (i_n - \bar{i}) (i_n - \bar{i})^H \right) \left( \sum_{k=1}^{N_b} (i_k - \bar{i}) (i_k - \bar{i})^H \right)^H \right\} \end{pmatrix}}
\]

\[
= \frac{1}{N_i - 1} \sqrt{\sum_{n=1}^{N_i} \sum_{k=1}^{N_b} (i_n - \bar{i})^H (i_n - \bar{i}) (i_n - \bar{i})^H (i_n - \bar{i})}
\]

\[
= \frac{1}{N_i - 1} \sqrt{\sum_{n=1}^{N_i} \sum_{k=1}^{N_b} (i_n - \bar{i})^H (i_n - \bar{i})^2}
\]

(5)

where \( i_n = I_n(x,y) \), \( \bar{i} = \bar{I}(x,y) \)

3.3.2 Offset: Forcing zero mean

While both examples in Algorithms 1 and 2 contained a parameter related to the offset, notice that Algorithm 3 does not. The main objective of \( \phi_n \) or \( \phi_n \) for the multi-band cases, as referred on the previous section, is to adjust each band of each image to the mean level the mean image \( \bar{I} \) has over that band. Obviously, it is necessary for a real homogenization that the mean levels are similar, and that irrelevant information, in LU/LC terms, is suppressed. In other words, it is necessary that an image taken at dawn or during the night, has the same average than another image captured at noon. However, the actual mean level each band has in the mean image \( \bar{I} \) carries no information at all. Taking this last fact into account, it makes much more sense to force each image to have zero mean and not waste computational effort in finding the optimal regression parameter \( \phi_n \). This approach is also inspired in the atmospheric corrections in [7], which just by subtracting a convenient value on each image, compensated the disturbances produced by the atmospheric phenomenon known as haze.

Depending on which application the data is targeted to, different post-processing techniques can be convenient to compensate for the mean subtraction. For classification purposes, it is most convenient
to leave the data as it is. The only reason why Algorithm 3 adds up \( \mu \) to every pixel in each image is an intent to bring the data back to its previous range, in order to ease its codification in 8 bit integers.

### 3.3.3 Projecting to the mean and iterative behavior

All previous approaches to RRN chose, from the set \( \{ \mathbf{I}_n \in \mathbb{R}^{N_y \times N_x \times N_b} \} \) of images they worked upon, one control image. Then, they attempted to bring all other images, through their algorithms, to the same space the control image was.

In this thesis’ proposal, Algorithm 3, all images are brought towards the mean image. The insight behind this fact is that there should not be any prioritisation among the images, and that the mean of all time realizations represents better how an area actually is than one arbitrarily chosen image. Intuitively, the mean of all the images after one iteration of Algorithm 3 will be an even more representative sample of how the region is, and so on with each iteration. Both this intuition and the possibility of combining this RRN algorithm with non-linear outlier detection methods to mitigate the effect of clouds and other disturbances (see Chapter 4) motivated the iterative behavior of Algorithm 3.

### 3.3.4 Further extensions

This algorithm is an entirely new proposal itself, and therefore, many issues and smaller studies remain that could have been done. The following presents the two most immediate ones.

Recall the first precondition on section 3.1, this is, \( \mathbf{I}_n \in \mathbb{R}^{N_y \times N_x \times N_b} \). On view of Algorithm 3, it is easy to see that there are two distinct points at which the precondition is required, the estimation of the mean on line 13, and the estimation of the covariance matrices of each pixel in line 15. In both cases a weight, \( \frac{1}{N_i} \) and \( \frac{1}{N_i - 1} \) respectively, is assigned to each image. This observation suggests a simple proposal for the case in which each image has a different number of bands, this is, \( \mathbf{I}_n \in \mathbb{R}^{N_y \times N_x \times N_{b_n}} \). All images could easily be extended with zeros to the maximum number of bands \( N_b = \max_n \{ N_{b_n} \} \), i.e. \( \tilde{\mathbf{I}}_n \in \mathbb{R}^{N_y \times N_x \times N_b} \). Once that done, it would be fairly easy to modify both the estimation of \( \bar{\mathbf{I}} \) and \( \Sigma(x, y) \) in order to account for each band having a different number of samples \( N_{b_n} \).

Finally, the number of iterations in Algorithm 3, in its current state, has been arbitrarily set. In the experiments presented in Chapter 7 there is no analysis on the convergence or the convergence rate. Because of the changes that will be introduced in the algorithm in Chapter 4, an ad-hoc rule has been used, i.e. the number of iterations had to be around \( \frac{N_i}{2} \). It is clear that a study of the convergence would provide a better understanding of the problem, and the iteration by
3.3 EXTENSION AND IMPLEMENTATION ISSUES

iteration tracking of the performance would verify some design hypothesis that remain not proven.
Algorithm 3 Iterative multi-band variance comprehensive linear regression to the mean

\( I_n \in \mathbb{R}^{N_y \times N_x \times N_b}, \forall n \in (\mathbb{Z} \cap [1, N_i]) \) / {\( N_y: \) #(rows), \( N_x: \) #(columns), \( N_b: \) #(bands), \( N_i: \) #(images)}

2: \( \mu = 0, \rho = 0 / \{ 0 \in \mathbb{R}^{N_b} \} \)
   
   for \( n = 1 \) to \( N_i \) do
   
   4: \( \rho = \frac{1}{N_x N_y} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} I_n(x, y) \)
   
   6: for all \( x, y \) do
   
   7: \( I_n(x, y) = I_n(x, y) - \rho \)
   
   8: end for
   
   end for

10: \( \mu = \frac{\mu}{N_i} \)

12: for iteration = 1 to \( It_{\text{max}} \) do

14: for all \( x, y \) do

15: \( \Sigma(x, y) = \left\| \frac{1}{N_i} \sum_{n=1}^{N_i} \left( [I_n(x, y) - \bar{I}(x, y)] [I_n(x, y) - \bar{I}(x, y)]^T \right) \right\|_F \)

16: \( \Omega(x, y) = \frac{1}{\Sigma(x, y)} \)

18: for \( n = 1 \) to \( N_i \) do

19: \( \hat{\alpha}_n = \text{argmin} \left\{ \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} \left( \Omega(x, y) \| \text{diag}(\hat{\alpha}_n) I_n(x, y) - \bar{I}(x, y) \|^2 \right) \right\} \)

20: for all \( x, y \) do

21: \( I_n(x, y) = \text{diag}(\hat{\alpha}_n) I_n(x, y) \)

22: end for

24: end for

26: for \( n = 1 \) to \( N_i \) do

27: for all \( x, y \) do

28: \( I_n(x, y) = I_n(x, y) + \mu \)

29: end for

30: \( \text{Where } \text{diag}(\hat{\alpha}_n) = \begin{bmatrix} \alpha_{n,1} & 0 & \ldots & 0 \\ 0 & \alpha_{n,2} & \ldots & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \ldots & \alpha_{n, N_b} \end{bmatrix} \)}
NON-LINEAR OUTLIER DETECTION AND REMOVAL

One of the main reasons why it is uncommon to see RRN studies with as many images as in this thesis (see Section 1.2.2), is because of the difficulty in finding and choosing them. Disturbances such as clouds or reflections over the sea can have a strong effect on the algorithm’s results, and definitely perturb LU/LC classifiers. This study is no exception, and therefore quite a large number of the images included in this thesis contain a relevant level of contamination by clouds, haze, and other phenomena. As it has been pointed out in Section 3.2, however, this thesis’ proposal for RRN is practically immune to the effects of such disturbances. Nonetheless, these disturbances do remain in the final corrected images, still making their LU/LC classification hard to solve.

Section 4.1 introduces the theoretical reasoning used to propose Algorithm 4 in Section 4.2, an ad-hoc approach designed to eliminate or at least soften the effect of these disturbances on LU/LC results. As it can be seen, this Algorithm will be embedded inside Algorithm 3 at the end of each iteration.

4.1 THEORETICAL BACKGROUND

4.1.1 Rician modeling

The following will suggest that, if both illumination effects and variation within a class can be modeled as Gaussian processes, and they interact additively, an appropriate model for the Frobenius norm of the covariance matrix estimation is the Rician one.

Assume $N_i$ images that will be indicated by the index $t$. Suppose the effect of the illumination on any random pixel in image $t$ can be modeled as an additive Gaussian noise $c_{ill}(t) \sim N(\mu_{ill}(t), C_{ill}(t))$. Assume therefore that the value of a pixel that falls within one of the classes $k = 1, \ldots, N_{cl}$, can be expressed as $z_{k \in A_k(t)} = c_k(t) + c_{ill}(t)$, where $A_k(t)$ is the set of points inside class $k$ in image $t$, and $c_k(t) \sim N(\mu_k(t), C_k(t))$. The probability density function (PDF) for any pixel of class $k$ in image $t$, $f_{z(t)}(z(t) \in A_k(t))(z(t))$ will be Gaussian, and specifically the Gaussian resulting from the convolution of the previously
stated distributions. This is mathematically expressed in Equation (6). Now, if the union between the sets of all classes covers the whole image, i.e. the classes describe the image completely, the PDF for a random pixel in the image can be derived. Applying the law of total probability it is proven in Equation (7) that this PDF, \( f_{z(t)}(z(t)) \), will also be Gaussian.

\[
f_{z(t)|z(t)\in A_k(t)}(z(t)) = N(\mu_{ill}(t), C_{ill}(t)) * N(\mu_k(t), C_k(t)) \tag{6}
\]

\[
f_{z(t)}(z(t)) = \sum_{k=1}^{N_i} \left( Pr \{ z(t) \in A_k(t) \} f_{z(t)|z(t)\in A_k(t)}(z(t)) \right) \tag{7}
\]

Recall the final simple expression of the Frobenius norm of the covariance matrix estimation in Equation (5). Using the newly defined notation it can be expressed as in Equation (8). Note that \( \sigma \) is exactly \( \Sigma(x,y) \), but where the choice of pixel \( x,y \) is done randomly. Each term \( z^H(t)z(\tau) \) follows a generalized chi-squared distribution, because any product of two Gaussian variables can be expressed as a chi-squared distribution and the sum of chi-squared variables is chi-squared distributed. Pitifully, deriving the exact PDF for \( \sigma \) is well out of this thesis' scope. However, [31] derives the expression for the PDF of the product of non-central chi-squared distributions, which particularizes in the case at hand as Equation (9), where \( \psi = [z^H(t)z(\tau)]^2 \) and \( k(t,\tau) \) is the degree of freedom of the chi-squared distribution that regulates \( z^H(t)z(\tau) \).

\[
\sigma = \frac{1}{N_i - 1} \sqrt{\sum_{t=1}^{N_i} \sum_{\tau=1}^{N_i} [z^H(t)z(\tau)]^2} \tag{8}
\]

\[
f_\psi(\psi) = \frac{\psi^{k(t,\tau) - 1}K_0(\psi^{1/2})}{2^{k(t,\tau) - 1} \left[ \Gamma \left( \frac{k(t,\tau)}{2} \right) \right]^{1/2}} \tag{9}
\]

Summarizing, \( \sigma \) is the sum of \( N_i^2 \) random variables distributed as stated in Equation (6). Considering the structure of the distribution in Equation (9) is similar to that of the chi-squared distribution, all available distributions of the chi-squared family were considered and their match with the data was visually assessed for one experimental case. The most convincing match was observed with the Rician distribution, that from then on was assumed on \( \sigma \).

\[
f_\psi(\psi) = \frac{\psi^{1/2} - 1 e^{-\frac{\psi}{2}}}{2^{1/2} \Gamma \left( \frac{k(t,\tau)}{2} \right)} \tag{10}
\]

This assumption is made twice during Algorithm 4. At the beginning to make a hard decision on whether there are outliers or not in the

---

1 See the definition of the chi-squared distribution PDF in Equation (10).
4.2 Proposal

Algorithm 4 uses both the model reasoned in Section 4.1.1 and the intuitive interpretation of the variation maps given in 4.1.2 to detect and eliminate outliers.

Once a pixel \((x, y)\) in a specific image \(n\) is considered an outlier, it is included in the set \(B_n\). The belief in the fact that a position contains considered image set (see line 4). And once again at the end to establish which pixels will be considered outlier candidates and therefore modified (see line 16).

4.1.2 Variances for outlier detection

The following motivates the basic principles used in Algorithm 4 to detect outliers in the set of images \(I_n \in \mathbb{R}^{N_y \times N_x \times N_b}\). Outliers can be present in any pixel position \((x, y)\) at any point in time \(n\). The only known information about the outliers in this context are the following two statements,

- They have values that are far from the center of the data distributions,
- They are extremely uncommon.

It follows from this two properties that the variation map \(\Sigma(x, y)\) computed in Algorithm 3 might have higher values precisely in the positions \((x, y)\) that contain outliers at a specific instant \(n\). This would not necessary be the case, as it depends on exactly how far the typical outlier is from the nearest datapoint, and on how often in time the same point \((x, y)\) has an outlier. Luckily, it has been experimentally verified that this is so.

Assuming this as a fact, it is easy to deduce what the computations between line 6 and 8 in Algorithm 4 intend. By extracting one of the time points at a time, and recomputing the variation map in \(\Sigma_n\), one obtains, for each pixel \((x, y)\), \(N_i\) different values for the Frobenius norm of the covariance matrix. If there is an outlier in the pixel \((x, y)\), the lowest of these values will indicate in which time instant \(n\) is most likely to be. Finally, the variance within the variation maps values for each pixel \((x, y)\) will give a consistent indicator on wether the particular pixel does contain outliers.

Line 13 combines both previous intuitions to give a specific value to the belief that each position contains at least an outlier. Notice that, in consideration with the treatment a pixel gets once it has been pinpointed as outlier, this belief combination is done through a product. This way, only if a pixel has both a high \(\Sigma(x, y)\) and a high variation within the \(\Sigma_n\)s, it will be affected greatly by the outlier removal step in line 24.
an outlier, $\Theta(x,y)$, is normalized within $B_n$ so it ranges from zero to one. Then, it is used to control the proportion between the old value and the temporal mean the new pixel value will contain.

Note that in the Algorithm, a thresholding with the value that leaves the higher 15.87% of the distribution is done. This specific percentage was chosen because its parallelism with the Gaussian case, in which Equation (11) holds.

$$\int_{\mu+\sigma}^{+\infty} G_{\mu,c}(z) dz = 15.87\%$$  \hspace{1cm} (11)

**Algorithm 4 Model-based non-linear outlier detection and removal**

Line 23 in Algorithm 3

2: $f_\sigma(\sigma) = \text{fitdist}(\Sigma, 'Rician')$

Find $z$ such as: \( \int_{z}^{+\infty} f_\sigma(\sigma) d\sigma = 15.87\% \)

4: if $\#_{\Sigma(x,y)\geq z} \frac{N_xN_y}{N_xN_y} > 15.87\%$ then

6: for $n = 1$ to $N_i$ do

7: $\Sigma_n(x,y) = \left\lVert \frac{1}{N_i-2} \sum_{k\neq n} (I_k(x,y) - I(x,y))^2 \right\rVert_F$

8: end for

10: for all $x,y$ do

12: $m = \frac{1}{N_i} \sum_{n=1}^{N_i} \Sigma_n(x,y)$

13: $\Theta(x,y) = \Sigma(x,y) \sqrt{\frac{1}{N_i-1} \sum_{n=1}^{N_i} (\Sigma_n(x,y) - m)^2}$

14: end for

16: $B_n = \{(x,y) : (\Psi(x,y) = n) \& (\Sigma(x,y) > z)\}$

18: $p = \max_{(x,y)\in B_n} (\Theta(x,y))$

20: $q = \min_{(x,y)\in B_n} (\Theta(x,y))$

22: for all $(x,y) / (\Psi(x,y) = n)$ do

23: $\Theta(x,y) = \frac{\Theta(x,y) - q}{p-q}$

25: end for

24: $I_{\Psi(x,y)}(x,y) = (1 - \Theta(x,y)) I_{\Psi(x,y)}(x,y) + \Theta(x,y) I(x,y)$

26: end if

Line 24 in Algorithm 3

38
Part III

CLASSIFICATION AND SEGMENTATION
Once the homogenization algorithm is proposed, its testbench has to be decided. In Section 2.1 it was clarified that crossed-classification would be the selected method. However, this still leaves the decision on which classifier is going to be used. In order to decide this, that will condition all the experimental results in the thesis, a procedure both based in the current literature and on experimental results was used. This procedure will be explained and detailed in this Chapter.

Section 5.1 specifies the data selected to obtain the experimental results and Section 5.2 exposes which classifiers were tested, the reasoning behind it, and the obtained results.

5.1 REDUCED DATASET SELECTION

First, the dataset described in Table 2 was selected. Notice that the images were pre-processed by applying the 6S radiative transfer code \(^1\) [17] [16], a commonly used RN algorithm. Observe also that images from three different satellites are present, and that the dates in which they were captured ranges from 1984 to 2010. Notice also, that contrarily to what is done in Chapter 7, all classification results given in this Chapter will be self-classification results. Moreover, each classification will use all the available bands, and occasionally some extra features.

As the procedure performed in each of them was rather extensive, i.e. it included visual assessment of every classification result, intensive experimentation was only done on images B and C. Both images were selected from the region of Gran Canaria, because its steep orography and intertwined land classes make it a much harder area for automatic LU/LC classification. Images A, D, E and F were reserved for verification of the classifier quality after the actual selection.

---

1 As opposed to crossed-classification results. The test points and train points will be from the same image.
5.2 CLASSIFIER SELECTION

Table 2: Description of the dataset used to obtain experimental background for the selection of a classifier. All images were radiometrically corrected by the 6S radiative transfer code \[17,16\] before experimentation. Refer to Table 1 for specific information on each satellite’s properties.

<table>
<thead>
<tr>
<th>Name</th>
<th>Location</th>
<th>Satellite</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image A</td>
<td>Gran Canaria</td>
<td>LANDSAT 5</td>
<td>9th of July 2009</td>
</tr>
<tr>
<td>Image B</td>
<td></td>
<td>LANDSAT 5</td>
<td>22nd of September 1984</td>
</tr>
<tr>
<td>Image C</td>
<td></td>
<td>SPOT 5</td>
<td>1st of August 2009</td>
</tr>
<tr>
<td>Image D</td>
<td></td>
<td>SPOT 1</td>
<td>25th of March 1988</td>
</tr>
<tr>
<td>Image E</td>
<td>Senegal</td>
<td>LANDSAT 5</td>
<td>26th of January 2010</td>
</tr>
<tr>
<td>Image F</td>
<td></td>
<td>LANDSAT 5</td>
<td>21st of January 1985</td>
</tr>
</tbody>
</table>

5.2 CLASSIFIER SELECTION

The six different tested classifiers were Support Vector Machines (SVM), Neural Networks (NN), Minimum Mahalanobis Distance, Maximum Likelihood, Minimum Distance and Parallelepipeds. These are the classifiers that are most commonly found in LU/LC applications in the remote sensing literature. Further insight on the principles behind each of them or their application to LU/LC classification can be found in [29].

5.2.1 Parametric classifiers, parameter selection

Both SVM and NN classifiers need parameter calibration. Contrarily to the results in Chapter 7 however, the tests presented here were implemented inside a private remote sensing software framework, ENVI version 5.0 (Exelis Visual Information Solutions, Boulder, Colorado). This framework proposes default parameters for every image and parametric classifier selected.

For the NN case, in which the literature on selecting the proper parameters [2] tends to agree in that every case requires a specific heuristic process, the default parameters from ENVI were used. The only constraint that was given was for the number of hidden layers to be one.

In the SVN case, on the other hand, the recent study [15] on the choice of kernel functions when using SVN classifiers for LU/LC tasks was considered. This study suggests that RBF kernels are more suited to the task, and that their optimal parameters \((C, \gamma)\) could be efficiently found by an exponential grid search. These parameters represent, respectively, a control over the regularization of the convex optimization problem behind training SVN classifiers, and the width of the exponential kernel function. Intuitively, the higher \(C\) is, the more the resulting SVN is allowed to miss-classify points in its own
Figure 1.: Search on an exponential grid \(2^n\) of the optimal parameters for the RBF kernel \((C, \gamma)\) on image C before any preprocessing. The choice of \(\gamma\) was found irrelevant, and the optimal \(C\) is 1024.

The exponential grid search for the SVN RBF parameters can be observed in Figure 1. The training and test databases were sets of points within the area of Gran Canaria that have not changed their LU/LC classification during the considered time period.

5.2.2 Features

At the time of the experimentation it was considered to add extra features to all LU/LC classifications, if they were to improve classification results. According to the results that will be presented here (see Table 3), and the computational cost associated with it, the idea was finally discarded. The results are however still meaningful, and will, therefore, be included.

The features considered were selected according to two different parameters. First, it was necessary for them to be properly defined in all the considered dataset (see Table 2), i.e. not depend on bands not present in some of the satellites. Second, they needed to be commonly used or mentioned in the state of the art.
5.2 CLASSIFIER SELECTION

With these criteria, the selected features were the Normalized Difference Vegetation Index (NDVI, see Equation (13)), the Modification of the Normalized Difference Water Index (MNDWI, see Equation (14)), and the mean bands’ entropy within a 330 m × 330 m window. Notice that it is not possible to calculate the MNDWI for Image D, as the SPOT 1 satellite does not have a SWIR band. In these cases it is common to compute the Normalized Difference Water Index (NDWI, see Equation (15)) instead. Recall that NIR refers to Near Infrared and SWIR to Short Wave Infrared. Notice that the SWIR band can be found in the literature referred by the name of Middle Infrared (MIR). As for the last feature, the mean bands’ entropy, it is computed by calculating, for each band, the entropy within the specified window, and then averaging these entropy values for all bands.

\[
\text{NDVI} = \frac{\rho_\text{NIR} - \rho_\text{R}}{\rho_\text{NIR} + \rho_\text{R}} \tag{13}
\]

where \(\rho_\text{NIR}\) represents the NIR band value and \(\rho_\text{R}\) represents the Red band value.

\[
\text{MNDWI} = \frac{\rho_\text{G} - \rho_\text{SWIR}}{\rho_\text{G} + \rho_\text{SWIR}} \tag{14}
\]

where \(\rho_\text{G}\) represents the Green band value and \(\rho_\text{SWIR}\) represents the SWIR band value.

\[
\text{NDWI} = \frac{\rho_\text{G} - \rho_\text{NIR}}{\rho_\text{G} + \rho_\text{NIR}} \tag{15}
\]

where \(\rho_\text{G}\) represents the Green band value and \(\rho_\text{NIR}\) represents the NIR band value.

5.2.3 Results

The results of the different tests are shown in Table 3. There it can be seen that SVM based classifiers tend to give better performances in LU/LC applications than any other of the considered classifiers. The NN based ones have a tendency to be in second place but can get extremely harmed by the addition of features that do not carry new information.

Another conclusion of this study is that the information carried by the NDVI and MNDWI indices can already be extracted by non-linear classifiers directly from the spectrum. Therefore, adding any of these indices as a feature only influences the results by the curse of dimensionality. However, the drop in performance that occurs with the NN classifier when the NDVI or MNDWI are used cannot be explained only by this fact. This results suggests that, even if each
5.2 CLASSIFIER SELECTION

Table 3.: Accuracy over the different databases for each combination image - classifier - feature. In all tests the classifiers were trained and evaluated on the spectrum and the feature after the + symbol, when applicable. Results on the training database are included for completion.

of the indices provides useful information to separate the class it is designed for, it also introduces confusion on how to properly separate the other classes.

Finally, the accuracy improvements the mean bands’ entropy provides reveal that textural information can greatly improve LU/LC classification results, in line with the findings in [3]. The best obtained combination, which performed consistently across the images, was the SVM RBF classifier with any \( \gamma \) and \( C = 1024 \), using the spectrum and the Mean Entropy as features. When working directly on the spectrum, the same SVM RBF classifier performed better than any other, and therefore it will be used in all experiments in Chapter 7.
SEGMENTATION

The following Chapter will describe this thesis’ proposal to use spatial and textural information to improve sea / land classification accuracy. The intention behind this is to exemplify how spatial information is fundamental to LU/LC classification, and should be incorporated to improve current pixel by pixel classification algorithms. This Chapter, therefore, does not intend to present novel or ambitious techniques, but just to prove a concept.

The selected methodology will be to apply a segmentation algorithm, detailed in Section 6.2, to textural features derived from the original data, that will be presented in Section 6.1. Section 6.3 will deal on how to take profit of the $N_i$ different segmentations obtained from the previous to generate a single enhanced segmentation, and on how to mix the final results with the classification results obtained by LU/LC classification algorithms.

6.1 TEXTURE INFORMATION

As the results in Chapter 5 suggest, textural information generally improves LU/LC results greatly. During the actual testing performed in Chapter 5, it was also observed, however, that they mostly improved land / sea classification. This is because generally, sea areas have much less texture or variation than inland areas.

To take advantage of this fact, the feature described by Equation (16) was defined. According to what has been said, each term of the sum should provide a similar gray-scale image. The land area should appear brighter, because higher variations imply higher gradients, while the sea should appear dimmer. Note that the only reason the gradient magnitude is not used in the computation of $F(I_n, \sigma)$ is to speed up calculations through the use of the approximation in Equation (17).

This resulting image adapts specially well to a well-known segmentation method, part of a wide family of segmentation algorithms referred to as snakes or Active Contours. The specific method is known as the Chan & Vese algorithm for segmentation, and it ill be presented in the next section.
The segmentation algorithm used for this purpose is available both in a scalar and a vector version \[6, 5\]. Explaining or giving an understanding on the exact theoretical principles behind it is out of the scope of this thesis. A basic understanding on what does the scalar version do will be given here from an input / output point of view. Later, the final procedure used in this thesis to perform a land / sea segmentation will be presented.

The vectorized version \[5\] was tested but discarded for two different reasons. Primarily, the scalar version \[6\] gave satisfactory enough results when working on the textural feature defined in Equation (16). Secondly, the image size was already bigger than it is usual for segmentation problems, causing the progress to be very slow. Adding more dimensions to the problem would have made the experimentation costs in memory and time too high.

The original article for the scalar version of the Chan & Vese for segmentation \[6\] describes the cost function the algorithm is trying to minimize as Equation (18), where \(F_n\) is the input image, \(F(I_n, \sigma)\) in this thesis’ case. Summarizing, the Chan & Vese algorithm minimizes \(\Gamma\) and finds the two regions within the image that define, in a MSE sense, the closer to two different constant regions. The parameters for this algorithm were not optimized for the case at hand, and where left to their defaults specified in \[6\], i.e. \(\mu = 0.2, \lambda_1 = \lambda_2 = 1\) and \(\nu = 0\).

The Chan & Vese algorithm is solved iteratively, and it does not guarantee to converge to the optimal solution. It also requires an initial curve \(C_0\), that will greatly influence how the final solution will be. Note that the curves \(C\) and \(C_0\) can be in fact several curves, but all of them need to be closed, clearly separating \text{inside}(C)\) and \text{outside}(C) in two distinct sets that cover the whole scene. In the case at hand, the initial curve \(C_0\) is obtained, for each image \(I_n\), through a thresholding of the corresponding \(F_n = F(I_n, \sigma)\). According to what was

\[
F(I_n, \sigma) = \sum_{z=1}^{N_{bn}} \left[ |\nabla_x \cdot (G_\sigma \ast I_n(z))| + |\nabla_y \cdot (G_\sigma \ast I_n(z))| \right]
\]

(16)

\[
N_{bn} : \text{Number of Bands in image } I_n
\]

\[
I_n : \text{Image } n
\]

\[
G_\sigma : \text{Gaussian kernel with variance } \sigma
\]

\[
\| \nabla \cdot (G_\sigma \ast I_n(z)) \| = \sqrt{ |\nabla_x \cdot (G_\sigma \ast I_n(z))|^2 + |\nabla_y \cdot (G_\sigma \ast I_n(z))|^2 }
\approx |\nabla_x \cdot (G_\sigma \ast I_n(z))| + |\nabla_y \cdot (G_\sigma \ast I_n(z))| \quad (17)
\]
previously established, the sea areas have lower values in $F_n$. Therefore, the thresholding is done on the value that has a $\text{Sea}_{\%}$ of the pixels in $F_n$ under it. $\text{Sea}_{\%}$ is practically a parameter and is supposed to be known in the studied region. After 4000 iterations, a curve $C_n$ is obtained, and the biggest closed connected region inside it is assumed to contain the land.

$$\Gamma (c_1, c_2, C) = \mu \| l(C) + \nu A(C)$$
$$+ \lambda_1 \int_{\text{inside}(C)} |F_n(x, y) - c_1|^2 \, dx \, dy$$
$$+ \lambda_2 \int_{\text{outside}(C)} |F_n(x, y) - c_1|^2 \, dx \, dy$$  \hspace{1cm} (18)$$

where $C$ is a curve, $l(C)$ its length and $A(C)$ its inside area.

### 6.3 Combining Results

#### 6.3.1 Combining multiple segmentations

After applying the steps described in Section 6.2, we have $N_i$ different segmentations of the same scene. Recall that each of them has a different resolution, depending on from which satellite the image $I_n$ came from. Therefore, a first step to mix all of them, will be to bring them all to the same resolution. On binary images such as segmentations, this is easily done by a nearest neighbor interpolation. In order to use all the possible information, all of them are transformed to the highest resolution, i.e. to the highest number of pixels.

Once all of them are on the same resolution, a technique to generate a single enhanced segmentation is needed. [18] claims that when all classifications, or in this case, segmentations, have the same intrinsic value or quality, the majority voting technique is optimal. Intuitively, the segmentations that come from a higher resolution image should be more accurate than the others. In this case, however, higher resolution also means higher variations within the sea, as its texture starts to be distinguishable. Summarizing, lower resolution images generate segmentations that are rough in the edges and do not exactly follow the contour of the land, but higher resolution images generate segmentations that, while having most accurate edges, include more random mistakes. Therefore, as there is no intrinsic reason to believe one segmentation is better than another, the majority voting technique is used, i.e. each pixel is assigned to the class (land / sea) that most segmentations designated for it.
6.3.2 Combining segmentation and classification

In this case, and according to the objectives announced at the beginning of this Chapter, the enhanced segmentation information has only been used to compensate LU/LC classification errors, after this one was completed.

In other situations, however, this result could have been used in much bolder ways. It will be seen in Chapter 7 that the final segmentation has no Sea $\rightarrow$ Land errors. This means that all points labeled as Sea by the segmentation algorithm are indeed Sea. This repetitively tested property of the system above could, for example, be used to extend the training database the classifier uses for the class Sea, greatly improving classification results.
EXPERIMENTAL EVALUATION

This section presents the experimental results obtained during this thesis. All developed methods are assessed by empirical results based on ground truth data. In Section 7.1, different testbenches for Algorithm 3 are explained, and their results presented. In Section 7.2, the segmentation results of the technique proposed in Chapter 6 are presented and discussed.

7.1 HOMOGENIZATION - RRN

This section details how the RRN algorithm with outlier detection and removal capabilities has been tested. In other words, the results presented are obtained by combining Algorithms 3 and 4.

Section 7.1.1 specifies the preprocessing the images receive to fulfill the preconditions for Algorithm 3 (see Chapter 2). Section 7.1.2 compares the crossed classification results obtained by applying the 6S radiometric correction \[\{17, 16\}\] to a reduced dataset, with the results obtained by applying the methodology developed in this thesis. Section 7.1.3 presents the crossed classification results obtained by using all the available images from the region of Gran Canaria. On the other hand, Section 7.1.4 discusses the pro et contra of using only subsets of this dataset that were captured by satellites of the same family.

All LU/LC classification results in this section included the classes: Urban areas, Arid areas, Vegetation and Water. In every case, the classification accuracy is computed as the number of properly classified pixels over the number of pixels. This is done on a test database that includes the same number of pixels for each of the considered classes. The crossed classification results shown in Sections 7.1.2 to 7.1.4 are represented by color coded matrices. The same results but in a tabular manner can be found in the Appendix of this thesis.

7.1.1 Preprocessing, spatial resolution

Section 1.3 already announced that this thesis could not possibly correct all the diversity in remote sensing imagery. This implied: disregarding the differences between the bands definitions (see Table 1 in Chapter 2), using only the common bands to all the images available,
and more importantly, compensating for the different spatial resolutions.

There are two directions on which to operate in the scale space, i.e. the space of different resolutions of the same image. The resolution can be artificially improved, by interpolation techniques, going towards a more detailed scale. Or it can be degraded, going towards a worse one. Obviously, developing an interpolation technique that intelligently uses high resolution images of the same area to aid the interpolation of others would solve the issue of the different spatial resolutions. This task, however, is too broad and extensive to be done within this master’s thesis, and using current algorithms that do not take into account the variability of remote sensing imagery was considered too risky. Therefore, the degradation option has been chosen.

In each experiment in Sections 7.1.2 to 7.1.4, all images have been converted to the lowest resolution. To do that, the most common technique is the Gaussian scale space.

In a computer vision environment, the Gaussian scale space is used as a tool to analyze edges and patterns in an image at different resolutions, simulating the effect of seeing an object from a closer or further position. However, there is usually no need of establishing the dimension of the pixels, or how a particular Gaussian degradation corresponds to a particular change of resolution in the real world.

Specifically, the procedure followed here to decrease the resolution is the following. The change of scale $\gamma$, i.e. the ratio between the the desired dimension of the image and the actual dimension, is computed. Then, a Gaussian kernel with variance proportional to $\gamma$ is chosen and applied to the image. Finally, the image is consequently subsampled. More specifically, the variance of the Gaussian kernel has been chosen to be $2k\gamma$, and $k$ has been adjusted experimentally.

Table 4 describes the two images used to adjust $k$. As it can be seen in Table 1 from Chapter 2, the two satellites involved have different resolutions. Moreover, the selected images were captured on really close dates and therefore it can be assumed that their LU/LC classification did not change. The false color representation of these images can be seen in Figure 2, which makes obvious that the characteristic heterogeneity of remote sensing imagery is present in this two image dataset too.

<table>
<thead>
<tr>
<th>Name</th>
<th>Satellite</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image 1</td>
<td>SPOT 2</td>
<td>29th of July 2009</td>
</tr>
<tr>
<td>Image 2</td>
<td>LANDSAT 5</td>
<td>20th of July 2009</td>
</tr>
</tbody>
</table>

Table 4: Details of the images involved in the proportionality constant choice. Their false color representations can be seen in Figure 2.
The specific procedure to choose the proportionality constant $k$ was to perform the convolution with a set of different kernels with parameter $k$ ranging from 0.2 to 4 and to evaluate the crossed LU/LC classification results. These results can be observed in Figure 3. There, $n \rightarrow m$ refers to the crossed classification accuracy when training the classifier on image $n$ and testing it in image $m$. Taking into consideration these results, the proportionality constant was chosen $k = 0.5$. 
Figure 2.: Images involved in the proportionality constant choice. As it can be seen in their details in Table 4, these images were captured with a difference of only 9 days. Note that, while the LU/LC information practically does not change, the diversity within them is high, with clouds, atmospheric issues, and different sensors.
Crossed LU/LC classification accuracy with $k$

Figure 3: Experimental choice of the proportionality constant between the scale change and the Gaussian kernel standard deviation. The chosen is finally $k = 0.5$ according to crossed LU/LC classification results. The false color representation of the images involved can be seen in Figure 2 and their details found in Table 4.
7.1.2 6S radiometric correction as homogenization

Section 1.2.2 detailed that, other than Relative Radiometric Normalization algorithms, the literature also included Radiometric Normalization algorithms. These algorithms intend to recover the initial ground reflectance for each image, instead of bringing each image to a common ground statistically. However, these algorithms require ancillary information that can only be obtained through intensive work by experts.

As part of the dataset used in this thesis, there are 3 images that are available both in their original digital levels version and corrected through the 6S radiative transfer code. Even though the purpose of the 6S correction is not to homogenize the data, but to obtain the initial ground reflectance, this should theoretically provide more homogeneous data as well. The details of this images can be seen in Table 5.

<table>
<thead>
<tr>
<th>Name</th>
<th>Satellite</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image 1</td>
<td>LANDSAT 5</td>
<td>9th of July 2009</td>
</tr>
<tr>
<td>Image 2</td>
<td>SPOT 1</td>
<td>25th of March 1988</td>
</tr>
<tr>
<td>Image 3</td>
<td>SPOT 5</td>
<td>1st of August 2007</td>
</tr>
</tbody>
</table>

Table 5.: Details of the images available in both original and 6S corrected version.

As it has been established in Chapter 2, the metric to measure the obtained homogeneity across the images will be crossed classification results. In this thesis, crossed classification results will be always shown in the form of images, where the colormap is related to the accuracy and the axis are related to which image was used to train and test in each case.

Figure 4 shows four different crossed classification results. The top-left matrix (A) specifies the crossed classification accuracies for the original digital level version of the images. The top-right matrix (B) shows the results after applying the RRN methods developed in this thesis (Algorithms 3 and 4). The bottom-left results (C) show crossed classification results on the 6S corrected version of the images. Finally, the bottom-right matrix (D) shows the results when applying the developed methodology to the 6S version of the images.

These results suggest the following conclusions. First, Algorithm 3 seems to be accomplishing its task, resulting in a consistent improvement of the crossed classification results, seen by comparing A and B in Figure 4. Second, the 6S correction not only does not improve but worsens crossed classification accuracies. On the other hand, while Algorithm 4 does not seem to improve self classification results by

---

1 Referring to classification of one image using training points from the same image.
removing outliers, the 6S correction consistently improves self classification results. Interestingly, those crossed classification results that were greatly worsened by the 6S correction can not be recovered by the developed RRN algorithm. Other than in those combinations, the classification results in D are generally better than in any of the other combinations.

Notice, additionally, that this experiment shows that even with low quality estimates of both the mean and the variance at each pixel, due to the low number of samples, Algorithm 3 outperforms the 6S correction in crossed classification terms.
Figure 4.: Crossed LU/LC classification results on reduced dataset. Only images that were available both before and after the 6S radiometric correction [17, 16] were used, their details can be found in Table 5. A shows the results on the original images without preprocessing. B shows the results after the images were processed by Algorithms 3 and 4. C shows the results on the 6S corrected versions of the images. D shows the results after applying the developed algorithms to the 6S corrected versions of the images.
7.1.3 RRN on Full Gran Canaria dataset

This thesis strongly relies on the presence of the comparatively enormous (see Section 1.2.2) dataset provided by ULPGC and the TELECAN project on the island of Gran Canaria. The details of the images in the dataset can be read in Table 6.

<table>
<thead>
<tr>
<th>Name</th>
<th>Satellite</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image 1</td>
<td>LANDSAT 5</td>
<td>4\textsuperscript{th} of July 1984</td>
</tr>
<tr>
<td>Image 2</td>
<td></td>
<td>29\textsuperscript{th} of July 1987</td>
</tr>
<tr>
<td>Image 3</td>
<td></td>
<td>20\textsuperscript{th} of July 2007</td>
</tr>
<tr>
<td>Image 4</td>
<td></td>
<td>9\textsuperscript{th} of July 2009</td>
</tr>
<tr>
<td>Image 5</td>
<td></td>
<td>25\textsuperscript{th} of July 2009</td>
</tr>
<tr>
<td>Image 6</td>
<td></td>
<td>17\textsuperscript{th} of January 2010</td>
</tr>
<tr>
<td>Image 7</td>
<td>SPOT 1</td>
<td>25\textsuperscript{th} of March 1988</td>
</tr>
<tr>
<td>Image 8</td>
<td></td>
<td>24\textsuperscript{th} of July 2006</td>
</tr>
<tr>
<td>Image 9</td>
<td></td>
<td>1\textsuperscript{st} of January 2007</td>
</tr>
<tr>
<td>Image 10</td>
<td>SPOT 2</td>
<td>2\textsuperscript{nd} of July 2007</td>
</tr>
<tr>
<td>Image 11</td>
<td></td>
<td>29\textsuperscript{th} of July 2007</td>
</tr>
<tr>
<td>Image 12</td>
<td></td>
<td>1\textsuperscript{st} of January 2008</td>
</tr>
<tr>
<td>Image 13</td>
<td></td>
<td>13\textsuperscript{th} of January 2009</td>
</tr>
<tr>
<td>Image 14</td>
<td>SPOT 4</td>
<td>21\textsuperscript{st} of January 2006</td>
</tr>
<tr>
<td>Image 15</td>
<td></td>
<td>9\textsuperscript{th} of October 2006</td>
</tr>
<tr>
<td>Image 16</td>
<td></td>
<td>12\textsuperscript{th} of October 2007</td>
</tr>
<tr>
<td>Image 17</td>
<td></td>
<td>6\textsuperscript{th} of November 2010</td>
</tr>
<tr>
<td>Image 18</td>
<td>SPOT 5</td>
<td>27\textsuperscript{th} of September 2005</td>
</tr>
<tr>
<td>Image 19</td>
<td></td>
<td>4\textsuperscript{th} of August 2007</td>
</tr>
<tr>
<td>Image 20</td>
<td></td>
<td>1\textsuperscript{st} of August 2009</td>
</tr>
<tr>
<td>Image 21</td>
<td></td>
<td>25\textsuperscript{th} of February 2010</td>
</tr>
</tbody>
</table>

Table 6.: Details of the images in the full Gran Canaria dataset.

These images were all used in an experiment to assess the quality of the homogenization methodologies developed in this thesis. In Figure 5, the crossed classification accuracies before and after applying these algorithms are shown.

The consistency with which the resulting accuracy improves shows that Algorithm 3 is robust, supporting the arguments in its favor given in 3.2, i.e. that is a natural soft extension of the previous state of the art algorithms (see Section 1.2.2 for details). However, it also proofs that Algorithm 4 does neither increase nor decrease self classification results. This does not give any information on whether any of its parts are functioning properly, or on which of its steps could be done better.

It can also be observed in the results in Figure 5 that Images 10, 19 and 20 show an anomalous behavior. These images are shown in [2](http://www.telecan.eu)
their false color representation in Figure 6. It can be observed that all of them contain strong reflections in the sea areas. In fact, these three are the only images that contain sea reflections strong enough to be noticeable to the naked eye. Therefore, the presence of the comparatively low results of columns / rows 10, 19 and 20 prove the fact that the proposed ad-hoc Algorithm does not completely achieve its goal. While it has been observed during the experiments that the outliers detection part works reliably, the deletion part should probably be redesigned.

Finally, the red patterns in Figure 5 suggest that the SPOT and LANDSAT images obtain better crossed classification results with other images from the same satellite family. This is not a surprise, since it can be seen in the satellite specifications (see Table 1) that satellites from the same family tend to share the definition of their bands.

To give a more intuitive vision on how the classification results change after applying the RRN algorithm, the full classification of Image 16 when the classifier is trained on Image 9 has been computed both before and after the RRN procedures. These full classifications can be found in Figure 7. Note in Figure 5 that this particular classification (9 → 16) has been specially picked to show a case with relevant improvement.
Figure 5.: Crossed LU/LC classification results on the whole Gran Canaria dataset. Top, before RRN corrections. Bottom, after RRN by Algorithms 3 and 4. The details of the images involved can be seen in Table 6.
Figure 6.: False color representation of the images with least satisfactory results in the crossed LU/LC classification test. The crossed classification results can be found in 5.
Figure 7: LU/LC classification of Image 16 when the classifier is trained on Image 9, before and after RRN by Algorithms 3 and 4. See Table 6 for details on the images.
In the previous section it was observed that crossed classification results tend to be better within satellite families. Given this fact, the natural next step is to check whether homogenization itself could yield better results when only using satellites with the same bands’ definitions.

Figures 8 and 9 show the crossed classification accuracies between images of the same family both when the RRN procedure is applied to the whole dataset and when it is applied to only images coming from the same satellite family. Notice that the image numbering is the same as the one in the previous sections, and therefore, the images descriptions can be found in Table 6.

Even when the results do change under these situations, it is difficult to select which results are better. While using only imagery from a certain family seems to improve some of the crossed classification accuracies much more than using the whole dataset, this last option provides a more consistent improvement of the results. This suggests that, while the incoherences in the band definitions can not be compensated by the designed RRN procedure, it is still sensible to use the bigger dataset. The plurality of environments, weather conditions, and other variable factors found in the bigger dataset provides more useful information to the RRN algorithm. It is this extra information that makes the crossed classification results more homogeneous. From another angle, this can be seen as an overfitting phenomena, but at a dataset level. Keeping only a particular subset of the dataset may improve particular results over the same subset but imply a loss of generality, which can hurt the homogeneity with which they are yield.
Results on SPOT images with global experiment

Results of only using the SPOT images

Figure 8.: Crossed LU/LC classification results in the SPOT images from the Gran Canaria dataset. Top, the results when Algorithms 3 and 4 are fed with all the images in the Gran Canaria dataset, from all satellites. Bottom, when the RRN is done using exclusively information from SPOT images. The image indices are the same that were used in Section 7.1.3 The image details can be found in Table 6.
Figure 9.: Crossed LU/LC classification results in the LANDSAT images from the Gran Canaria dataset. Top, the results when Algorithms 3 and 4 are fed with all the images in the Gran Canaria dataset, from all satellites. Bottom, when the RRN is done using exclusively information from LANDSAT images. The image indices are the same that were used in Section 7.1.3. The image details can be found in Table 6.
7.2 SEGMENTATION

The segmentation proposal can be evaluated in two different aspects. First, the results of the Chan & Vese algorithm on the textural feature presented in this thesis can be assessed. Later, the mixing procedure can be evaluated. To do both steps ground truth data is needed. The Spanish government offers digital elevation models of every region in the country. The digital elevation model for the region of Gran Canaria and the ground truth data for segmentation derived from it can be seen in figure 10. As it can be observed, not all the island is covered by the ground truth data. Consequently, the following accuracy results will only refer to the region on which data is available.

This ground truth data has been used to evaluate the segmentation results in each image, before and after the mixing is performed. These results, which are considered to be quite accurate, can be seen in Figure 11. Importantly for the desired application of aiding LU/LC classification, it can be seen in the top plot of Figure 11 that no Land point is considered to be Sea after the mixing is performed. Figure 12 shows how the resulting segmentation fits one of the images in the dataset, and also a LU/LC classification that could be clearly improved using this segmentation. This could be done in two different ways. First, the classification of the sea could directly be modified according to the segmentation. Second, the results of the segmentation could be fed into the classifier as training database, ensuring that the class *Sea* is properly learned, which would eliminate the inland water false detections.
Figure 10: Top, Elevation model used to find ground truth data for segmentation, bottom.
Figure 11.: Segmentation error in the 21 images from the Gran Canaria dataset. The image details can be found in Table 6.
Top, percentage of land pixels segmented as sea. Middle, percentage of sea pixels segmented as land. Bottom, mean error percentage.
Figure 12.: Top, segmentation results superimposed in a false color representation of Gran Canaria. Bottom, classification of Image 16 when training on image 10, after the RRN procedure. The images’ details can be found in Table 6. It is clear that in this case, the segmentation results could aid LU/LC classification.
CONCLUSIONS

During this thesis, a novel approach to RRN has been proposed. It has been shown that it homogenizes the data as expected, performing consistently within an extremely broad set of conditions. Moreover, this algorithm has been compared, in terms of crossed classification, to algorithms that are currently being used for RN, i.e. to correct remote sensing imagery and obtain the ground reflectance. Finally, this algorithm has also been theoretically contextualized and it has been shown that is a logical evolution of the state of the art in RRN.

In a more theoretical diversion, crossed classification results have been theoretically assessed as a tool to measure the homogeneity between datasets of the same origin.

In a more practical diversion, an ad-hoc algorithm to detect and eliminate large typical disturbances has also been proposed. However, it has been found that its effects are not relevant enough to improve LU/LC crossed classification relevantly.

Finally, a proof of concept on using textural and spatial information to aid LU/LC classification has been done by designing, developing and experimentally verifying a segmentation algorithm with temporal information fusion capabilities.

In short, this thesis has accomplished all of its goals, additionally providing empirical basis to encourage further research both in the field of RRN and in the field of LU/LC classification with spatial and textural features.


Part V

APPENDIX
RESULTS

The following shows the quantitative version of the results that have been shown on Chapter 7 as color coded matrices. Due to space constraints, all results have been reported as integer percentages.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>71</td>
<td>66</td>
<td>53</td>
<td>60</td>
<td>49</td>
<td>60</td>
</tr>
<tr>
<td>2</td>
<td>48</td>
<td>71</td>
<td>68</td>
<td>78</td>
<td>75</td>
<td>63</td>
</tr>
<tr>
<td>3</td>
<td>56</td>
<td>72</td>
<td>81</td>
<td>76</td>
<td>71</td>
<td>68</td>
</tr>
<tr>
<td>4</td>
<td>60</td>
<td>70</td>
<td>73</td>
<td>79</td>
<td>76</td>
<td>65</td>
</tr>
<tr>
<td>5</td>
<td>60</td>
<td>71</td>
<td>72</td>
<td>78</td>
<td>76</td>
<td>65</td>
</tr>
<tr>
<td>6</td>
<td>64</td>
<td>58</td>
<td>62</td>
<td>67</td>
<td>71</td>
<td>84</td>
</tr>
</tbody>
</table>

Table 7.: Quantitative evaluation shown as a colored matrix in the bottom part of Figure 9. Corresponds to the crossed LU/LC classification accuracies for the LANDSAT images available. Data for the upper part of Figure 9 corresponds to the appropriate subset of Table 10.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>80</td>
<td>51</td>
<td>55</td>
</tr>
<tr>
<td>2</td>
<td>34</td>
<td>90</td>
<td>39</td>
</tr>
<tr>
<td>3</td>
<td>47</td>
<td>29</td>
<td>96</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>80</td>
<td>70</td>
</tr>
<tr>
<td>2</td>
<td>67</td>
<td>90</td>
<td>52</td>
</tr>
<tr>
<td>3</td>
<td>38</td>
<td>38</td>
<td>95</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>1</td>
<td>87</td>
<td>46</td>
</tr>
<tr>
<td>2</td>
<td>48</td>
<td>82</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>27</td>
<td>32</td>
<td>96</td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td>87</td>
<td>69</td>
</tr>
<tr>
<td>2</td>
<td>69</td>
<td>82</td>
<td>21</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>18</td>
<td>96</td>
</tr>
</tbody>
</table>

Table 8.: Percent crossed LU/LC classification accuracies shown as colored matrices in Figure 4. Experiments A, B, C and D are defined in the original Figure 4. Further details may be read in Section 7.1.2.
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>68</td>
<td>48</td>
<td>71</td>
<td>55</td>
<td>67</td>
<td>53</td>
<td>47</td>
<td>45</td>
<td>48</td>
<td>37</td>
<td>39</td>
<td>46</td>
<td>41</td>
<td>27</td>
<td>23</td>
<td>37</td>
<td>35</td>
<td>40</td>
<td>41</td>
<td>22</td>
<td>31</td>
</tr>
<tr>
<td>2</td>
<td>52</td>
<td>69</td>
<td>56</td>
<td>34</td>
<td>42</td>
<td>26</td>
<td>42</td>
<td>23</td>
<td>38</td>
<td>28</td>
<td>23</td>
<td>44</td>
<td>34</td>
<td>25</td>
<td>23</td>
<td>34</td>
<td>24</td>
<td>30</td>
<td>23</td>
<td>29</td>
<td>23</td>
</tr>
<tr>
<td>3</td>
<td>43</td>
<td>39</td>
<td>83</td>
<td>67</td>
<td>69</td>
<td>35</td>
<td>37</td>
<td>27</td>
<td>32</td>
<td>27</td>
<td>27</td>
<td>33</td>
<td>37</td>
<td>30</td>
<td>49</td>
<td>40</td>
<td>32</td>
<td>31</td>
<td>27</td>
<td>52</td>
<td>29</td>
</tr>
<tr>
<td>4</td>
<td>45</td>
<td>73</td>
<td>77</td>
<td>73</td>
<td>55</td>
<td>52</td>
<td>28</td>
<td>52</td>
<td>30</td>
<td>23</td>
<td>51</td>
<td>42</td>
<td>36</td>
<td>23</td>
<td>44</td>
<td>26</td>
<td>29</td>
<td>34</td>
<td>51</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>41</td>
<td>33</td>
<td>72</td>
<td>68</td>
<td>78</td>
<td>26</td>
<td>59</td>
<td>23</td>
<td>60</td>
<td>23</td>
<td>23</td>
<td>61</td>
<td>56</td>
<td>29</td>
<td>29</td>
<td>39</td>
<td>29</td>
<td>28</td>
<td>23</td>
<td>17</td>
<td>35</td>
</tr>
<tr>
<td>6</td>
<td>34</td>
<td>23</td>
<td>44</td>
<td>48</td>
<td>48</td>
<td>83</td>
<td>40</td>
<td>23</td>
<td>49</td>
<td>23</td>
<td>23</td>
<td>38</td>
<td>44</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>7</td>
<td>23</td>
<td>17</td>
<td>25</td>
<td>38</td>
<td>42</td>
<td>19</td>
<td>90</td>
<td>23</td>
<td>80</td>
<td>27</td>
<td>23</td>
<td>61</td>
<td>66</td>
<td>23</td>
<td>23</td>
<td>31</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>39</td>
</tr>
<tr>
<td>8</td>
<td>48</td>
<td>49</td>
<td>52</td>
<td>55</td>
<td>51</td>
<td>23</td>
<td>27</td>
<td>93</td>
<td>37</td>
<td>60</td>
<td>74</td>
<td>56</td>
<td>42</td>
<td>31</td>
<td>38</td>
<td>33</td>
<td>28</td>
<td>39</td>
<td>44</td>
<td>14</td>
<td>42</td>
</tr>
<tr>
<td>9</td>
<td>25</td>
<td>21</td>
<td>44</td>
<td>43</td>
<td>54</td>
<td>49</td>
<td>76</td>
<td>17</td>
<td>94</td>
<td>21</td>
<td>20</td>
<td>66</td>
<td>64</td>
<td>25</td>
<td>9</td>
<td>29</td>
<td>2</td>
<td>8</td>
<td>19</td>
<td>23</td>
<td>22</td>
</tr>
<tr>
<td>10</td>
<td>44</td>
<td>47</td>
<td>50</td>
<td>50</td>
<td>47</td>
<td>32</td>
<td>34</td>
<td>53</td>
<td>40</td>
<td>71</td>
<td>48</td>
<td>47</td>
<td>45</td>
<td>68</td>
<td>62</td>
<td>60</td>
<td>70</td>
<td>59</td>
<td>52</td>
<td>64</td>
<td>57</td>
</tr>
<tr>
<td>11</td>
<td>45</td>
<td>39</td>
<td>41</td>
<td>44</td>
<td>38</td>
<td>23</td>
<td>29</td>
<td>91</td>
<td>31</td>
<td>57</td>
<td>89</td>
<td>34</td>
<td>34</td>
<td>44</td>
<td>43</td>
<td>45</td>
<td>38</td>
<td>45</td>
<td>43</td>
<td>27</td>
<td>56</td>
</tr>
<tr>
<td>12</td>
<td>22</td>
<td>24</td>
<td>30</td>
<td>46</td>
<td>40</td>
<td>35</td>
<td>43</td>
<td>23</td>
<td>51</td>
<td>35</td>
<td>23</td>
<td>80</td>
<td>59</td>
<td>28</td>
<td>7</td>
<td>28</td>
<td>9</td>
<td>24</td>
<td>44</td>
<td>44</td>
<td>32</td>
</tr>
<tr>
<td>13</td>
<td>26</td>
<td>23</td>
<td>34</td>
<td>47</td>
<td>46</td>
<td>51</td>
<td>68</td>
<td>14</td>
<td>69</td>
<td>10</td>
<td>18</td>
<td>61</td>
<td>69</td>
<td>19</td>
<td>20</td>
<td>29</td>
<td>11</td>
<td>2</td>
<td>19</td>
<td>28</td>
<td>20</td>
</tr>
<tr>
<td>14</td>
<td>34</td>
<td>45</td>
<td>46</td>
<td>55</td>
<td>44</td>
<td>48</td>
<td>62</td>
<td>40</td>
<td>68</td>
<td>42</td>
<td>29</td>
<td>68</td>
<td>69</td>
<td>90</td>
<td>57</td>
<td>59</td>
<td>82</td>
<td>86</td>
<td>44</td>
<td>53</td>
<td>83</td>
</tr>
<tr>
<td>15</td>
<td>27</td>
<td>21</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>41</td>
<td>41</td>
<td>33</td>
<td>46</td>
<td>52</td>
<td>43</td>
<td>38</td>
<td>41</td>
<td>67</td>
<td>92</td>
<td>62</td>
<td>68</td>
<td>67</td>
<td>22</td>
<td>57</td>
<td>62</td>
</tr>
<tr>
<td>16</td>
<td>31</td>
<td>44</td>
<td>50</td>
<td>52</td>
<td>51</td>
<td>54</td>
<td>65</td>
<td>32</td>
<td>65</td>
<td>25</td>
<td>28</td>
<td>66</td>
<td>66</td>
<td>71</td>
<td>72</td>
<td>92</td>
<td>62</td>
<td>55</td>
<td>25</td>
<td>55</td>
<td>61</td>
</tr>
<tr>
<td>17</td>
<td>68</td>
<td>57</td>
<td>63</td>
<td>63</td>
<td>62</td>
<td>28</td>
<td>32</td>
<td>66</td>
<td>34</td>
<td>50</td>
<td>57</td>
<td>49</td>
<td>49</td>
<td>86</td>
<td>71</td>
<td>77</td>
<td>91</td>
<td>78</td>
<td>56</td>
<td>42</td>
<td>87</td>
</tr>
<tr>
<td>18</td>
<td>52</td>
<td>46</td>
<td>51</td>
<td>51</td>
<td>53</td>
<td>32</td>
<td>48</td>
<td>55</td>
<td>44</td>
<td>58</td>
<td>37</td>
<td>42</td>
<td>58</td>
<td>86</td>
<td>57</td>
<td>57</td>
<td>75</td>
<td>93</td>
<td>57</td>
<td>43</td>
<td>71</td>
</tr>
<tr>
<td>19</td>
<td>57</td>
<td>36</td>
<td>26</td>
<td>48</td>
<td>41</td>
<td>40</td>
<td>34</td>
<td>43</td>
<td>41</td>
<td>45</td>
<td>40</td>
<td>37</td>
<td>28</td>
<td>44</td>
<td>38</td>
<td>46</td>
<td>58</td>
<td>57</td>
<td>82</td>
<td>32</td>
<td>47</td>
</tr>
<tr>
<td>20</td>
<td>36</td>
<td>38</td>
<td>49</td>
<td>51</td>
<td>53</td>
<td>32</td>
<td>32</td>
<td>23</td>
<td>28</td>
<td>41</td>
<td>23</td>
<td>37</td>
<td>30</td>
<td>50</td>
<td>54</td>
<td>71</td>
<td>43</td>
<td>49</td>
<td>17</td>
<td>96</td>
<td>41</td>
</tr>
<tr>
<td>21</td>
<td>38</td>
<td>44</td>
<td>31</td>
<td>37</td>
<td>35</td>
<td>51</td>
<td>70</td>
<td>68</td>
<td>75</td>
<td>45</td>
<td>46</td>
<td>70</td>
<td>75</td>
<td>85</td>
<td>67</td>
<td>69</td>
<td>79</td>
<td>70</td>
<td>42</td>
<td>32</td>
<td>92</td>
</tr>
</tbody>
</table>

Table 9: Data represented in the upper part of Figure 5. Mean self classification accuracy: 84.4%. Mean crossed classification accuracy: 42.4%
Table 10: Data represented in the lower part of Figure 5. Mean self classification accuracy: 84.3%. Mean crossed classification accuracy: 62.8%
Table 11.: Quantitative evaluation shown as a colored matrix in the bottom part of Figure 8. Corresponds to the crossed LU/LC classification accuracies for the SPOT images available. Data for the upper part of Figure 8 corresponds to the appropriate subset of Table 10.