Automatizing photo interpretation of satellite imagery in the context of the Common Agriculture Policy subsidy control

Jonas Schmedtmann

Dissertação para a obtenção do Grau de Mestre em

Engenharia do Ambiente

Orientador: Doutor Manuel Lameiras de Figueiredo Campagnolo

Júri:

Presidente: Doutora Elizabeth Costa Neves Fernandes de Almeida Duarte, Professora Catedrática do Instituto Superior de Agronomia da Universidade de Lisboa.

Vogais: Doutor Manuel Lameiras de Figueiredo Campagnolo, Professor Associado, do Instituto Superior de Agronomia da Universidade de Lisboa;

Doutor Mário Sílvio Rochinha de Andrade Caetano, Professor Associado Convidado do Instituto Superior de Estatística e Gestão de Informação da Universidade Nova de Lisboa.

Lisboa, 2014

Acknowledgements

None of this would have been possible if it wasn't for the helpful guidance of my advisor Manuel Campagnolo. I am sincerely thankful for the many hours he devoted to this work, his patience, his enthusiasm. Research can be difficult, but he was encouraging when obstacles were found along the way. It was a pleasure working with him.

I would also like to thank Instituto de Financiamento da Agricultura e Pescas, especially Vitor Carmona and Manuel Simões from the control department, for giving me the opportunity to develop this work in a practical context. The real-world data they provided was the core element which made this work possible. Also their valuable feedback and insights to the world of subsidy control contributed to the success of my study.

Finally, this work is dedicated to my family and to my friends. Support, friendship and encouragement are just some of the words which best describe their contributions to all the time I spent at Instituto Superior de Agronomia. I want to thank every single one of them for five amazing years which made me grow in both a personal and a professional way. I wouldn't trade this experience for anything in the world.

Abstract

Computer Assisted Photo-Interpretation (CAPI) uses remotely sensed imagery to control farmers' subsidy applications in the context of the EU's Common Agriculture Policy. A simple and reproducible method to automatize CAPI in an operational context with the overreaching goal to reduce control costs and completion time was developed in this study. Validated control data provided by the Portuguese Control and Paying Agency for Agriculture (IFAP) and a multispectral atmospherically corrected Landsat ETM+ time series were used to calibrate and test the method. Taking advantage of the nature of subsidy declarations, object-based land cover classification for the 12 most controlled classes was carried out in the region of Ribatejo. The main feature of the presented method is that it allows choosing a confidence level on the automatic classification of farmers' parcels. While higher confidence levels reduce the risk of misclassifications, lower levels increase the number of automatic control decisions. A confidence level of 80% is a good compromise. This confidence level leads to over 55% of automatically taken control decisions with an overall accuracy of 84%. Furthermore, over 85% of all parcels classified as maize, rice, wheat or vineyard can be controlled by the method with the optimal confidence level.

Keywords: Common Agricultural Policy subsidy control, Landsat, Multitemporal analysis, Operational crop discrimination, Parcel-based classification, Remote sensing.

Resumo

A fotointerpretação assistida por computador (CAPI) utiliza imagens multiespectrais obtidas por detecção remota para controlar as candidaturas aos subsídios agrícolas no âmbito da Política Agrícola Comum. Neste estudo desenvolveu-se um método simples e reproduzível para automatizar o processo CAPI num contexto operacional, com o objectivo de reduzir custos e tempo de controlo. A calibração e o teste do método recorreram a dados de controlo obtidos e validados pelo Instituto de Financiamento da Agricultura e Pescas (IFAP) e a uma série temporal de imagens multiespectrais Landsat ETM+. A técnica de classificação baseada em objectos foi aplicada às 12 classes de ocupação do solo que estiveram mais sujeitas a controlo no ano de 2005 na região do Ribatejo. A principal característica do método apresentado prende-se com o facto deste permitir a escolha de um nível de confiança sobre a classificação automática das parcelas agrícolas. Com um nível de confiança de 80%, mais de 55% das decisões de controlo são tomadas automaticamente com uma precisão total de 84%. Para além disso, mais de 85% das parcelas classificadas como sendo milho, arroz, trigo ou vinha podem ser controladas pelo método com o nível de confiança óptimo.

Palavras-chave: Análise multi-temporal, Classificação baseada em objectos, Detecção remota, Discriminação de culturas em contexto operacional, Landsat, Política Agrícola Comum.

Resumo alargado

A Política Agrícola Comum (CAP) é, duma forma muito simples, um sistema de subsídios da União Europeia. De modo a ter direito aos subsídios agrícolas, os requerentes entregam no Instituto de Financiamento da Agricultura e Pescas (IFAP) uma candidatura onde inscrevem as parcelas da sua exploração e as ocupações culturais correspondentes, no caso português. Cerca de 5% das candidaturas são submetidas a "controlos de superfície" para garantir que os fundos comunitários sejam alocados de forma correcta. O controlo de uma dada candidatura pode ser feito por duas formas: visita ao campo ou fotointerpretação assistida por computador (CAPI), que recorre a imagens multiespectrais obtidas por detecção remota.

Neste estudo desenvolveu-se um método simples e reproduzível para automatizar o processo CAPI num contexto operacional, com a finalidade de reduzir custos e tempo de controlo, tendo em conta que o controlo de superfícies decorre num calendário bastante apertado. De forma simples, o objectivo do método é classificar a ocupação do solo de parcelas agrícolas desconhecidas de forma automática. Escolheu-se como área de estudo uma zona no Ribatejo, caracterizada por um clima mediterrâneo, onde se analisaram as 12 classes de ocupação do solo que estiveram mais representadas na amostra de controlo no ano de 2005. A calibração e o teste do método recorreram a uma série temporal de imagens multiespectrais Landsat ETM+ e a dados de controlo obtidos e validados no ano de 2005 pelo IFAP. Em detalhe, usaram-se 6 imagens multiespectrais com correcção atmosférica adquiridas entre Novembro de 2004 e Agosto de 2005 que melhor acompanharam os ciclos culturais das culturas mais importantes, bem como 11852 parcelas agrícolas com ocupação conhecida que cobrem cerca de 1057 km^2 da área em estudo.

O método apresentado fundamenta-se numa classificação baseada em "objectos", representados pelos contornos das parcelas declaradas. Esta técnica permite agregar os pixels contidos numa dada parcela agrícola e estimar desta forma a classe de ocupação do solo da mesma. A classe duma determinada parcela é estimada por um classificador através da assinatura espectral multi-temporal da parcela em questão, representada por um ponto no espaço das variáveis de classificação de dimensão 36 (6 imagens multiespectrais utilizadas com 6 bandas espectrais em cada uma). Foi realizada uma análise comparativa de dois classificadores que desempenham essa tarefa: k-Vizinhos mais Próximos (KNN) e Máquinas de Vectores de Suporte (SVM). Para além disso, aplicou-se um método de redução da dimensão do espaço das variáveis de classificação, com o objectivo de facilitar a análise e remover possíveis redundâncias nas imagens multiespectrais utilizadas.

A ideia que está por detrás do método desenvolvido é o facto de não ser viável automatizar a classificação de todas as parcelas agrícolas que requerem controlo, já que erros de classificação significativos seriam cometidos em algumas classes que são de discriminação mais difícil. Esta situação acontece quando as assinaturas espectrais multi-temporais de várias classes apresentam um comportamento muito semelhante. Tendo em conta a importância da precisão de um sistema de controlo no contexto operacional do IFAP, introduziu-se a noção de nível de confiança, que representa uma medida de precisão global a atingir em cada classe de ocupação do solo. Este

nível pode ser escolhido e ajustado pelo responsável do sistema de controlo. De um modo simples, níveis de confiança elevados reduzem o risco de erros de classificação, enquanto que níveis baixos aumentam o número de parcelas controladas de forma automática. Para determinar um nível de confiança óptimo, é preciso efectuar um balanço cuidadoso entre estes dois factores que são ambos desejáveis.

Conseguiu-se reduzir a dimensão do espaço das variáveis de 36 para 12 sem comprometer o desempenho dos classificadores de forma significativa, pelo que se considerou que a redução da dimensão do espaço é um procedimento útil. As variáveis mantidas dizem maioritariamente respeito às bandas nas regiões espectrais do vermelho e infravermelho próximo, resultado este que faz sentido do ponto de vista da detecção remota, já que estas duas bandas espectrais são particularmente úteis na identificação e discriminação de vegetação verde. Em relação à análise comparativa dos dois classificadores testados, o desempenho do classificador SVM em termos de precisão global de classificação foi superior ao apresentado pelo KNN.

No que diz respeito ao método de automatização do processo CAPI, um nível de confiança de 80% foi considerado como nível óptimo, garantindo pelo menos 80% de classificações correctas em cada classe de ocupação do solo. Com este nível de confiança, mais de 55% das decisões de controlo são tomadas automaticamente, com uma precisão total de 84%, que se considera como um resultado satisfatório. Para além disso, mais de 85% das parcelas classificadas como sendo “milho”, “arroz”, “trigo” ou “vinha” podem ser controladas pelo método com o nível de confiança óptimo. Por outro lado, nenhuma parcela nas classes “pousio”, “pastagem pobre” e “área não utilizada” consegue ser classificada de forma automática. Se um nível de confiança de 95% for exigido, apenas cerca de 20% das parcelas seriam controladas pelo método. Com este estudo demonstrou-se a viabilidade e o potencial do controlo automático em contexto operacional recorrendo a técnicas de detecção remota numa região mediterrânea.

Palavras-chave: Análise multi-temporal, Classificação baseada em objectos, Detecção remota, Discriminação de culturas em contexto operacional, Landsat, Política Agrícola Comum.

Contents

List of Figures

List of Tables

List of Acronyms

6S	Second Simulation of a Satellite Signal in the Solar Spectrum
ACP	Automatic Classification Proportion
BAR	Barley
CAP	Common Agricultural Policy
CAPI	Computer Assisted Photo-Interpretation
CwRS	Control with Remote Sensing
EC	European Commission
ETM+	Enhanced Thematic Mapper
EU	European Union
FAL	Fallow
FOR	Forage crops
HR	High Resolution
IFAP	Instituto de Financiamento da Agricultura e Pescas
JRC	Joint Research Centre
KNN	k-Nearest Neighbors
LEDAPS	Landsat Ecosystem Disturbance Adaptive Processing System
LPIS	Land Parcel Identification System
MAI	Maize
MARS	Monitoring Agricultural Resources
MODIS	Moderate Resolution Imaging Spectroradiometer
NCPA	National Control and Paying Agency
NDVI	Normalized Difference Vegetation Index
NIR	Near Infrared
NUA	Non used area
OA	Overall Accuracy

OAT Oat

OLI Olive grove

OTSC On-The-Spot Check

PA Producer's Accuracy

PGL Permanent grassland

POG Poor grassland

RBF Radial Basis Function

RIC Rice

SLC Scan Line Corrector

SVM Support Vector Machines

SWIR Short Wave Infrared

UA User's Accuracy

USGS United States Geological Survey

VHR Very High Resolution

VYA Vineyard

WHE Wheat

1 Introduction

1.1 Subsidy control in the context of the Common Agriculture Policy

The Common Agricultural Policy (CAP) is a system of European Union (EU) agricultural subsidies and programs that represents over 40% of the EU's budget, equivalent to €58 billion in 2011 (EC, 2013b; MARS, 2008). The CAP was introduced in 1962 and has changed over time depending on numerous factors such as developments of the sector and challenges it faces (Singh et al., 2014). Farmers have been able to receive direct subsidies based on cultivated area instead of production, the so called area-based subsidies, since the CAP reform in 1992. These direct payments are intended to compensate for farmers' reduced incomes due to the continuos decline of agricultural product prices. They help to provide a steady income for farmers, supporting the long-term viability of farms and protecting them against price fluctuations (Blaes et al., 2005; EC, 2013b). The CAP's expenditures cover: income support for farmers (direct payments, accounting for about 70%), rural development (measures to help farmers modernize their farms and become more competitive, 20%) and market support measures (10%) (EC, 2013b).

The main objectives of the CAP are to ensure a fair standard of living for farmers and to provide a stable and safe food supply at affordable prices for consumers. As for the latest CAP reform in 2013, the European Commission (EC) has proposed three priorities: viable food production, sustainable management of natural resources and balanced development of rural areas throughout the EU (EC, 2013b). Moreover, this reform introduces a set of so-called greening measures, which means that part of the direct payments will be conditional on the farmer using sustainable farming practices, thus improving the environmental performance of farming in Europe. This greening consists of agricultural practices beneficial for the climate and the environment: crop diversification, maintaining existing permanent grassland and having ecological focus area on the agricultural area (e.g. land left fallow, terraces, landscape features, buffer strips) (EC, 2013a,b).

Each member state has a National Control and Paying Agency (NCPA) which we call the "decision-maker", responsible for paying and controlling farmers, among other duties. In Portugal, these tasks are performed by Instituto de Financiamento da Agricultura e Pescas (IFAP). To apply for financial support, farmers are required to submit an application to their NCPA in the first six months of each year. This application includes the position and delineation of all their cultivated agricultural parcels. The exact area and crop type must be declared for each parcel. Before any payment is made, all applications are subjected to administrative checks. Furthermore, in order to ensure that CAP funds are spent appropriately, the responsible NCPA has to verify at least 5% of the applications by so-called On-The-Spot Check (OTSC). For area-based subsidies, this means that an agricultural parcel must be controlled at two different levels: both the declared crop and area must be correct (Blaes et al., 2005; MARS, 2008). In

order to facilitate the distribution of the CAP subsidies, each NCPA runs an Integrated Administration and Control System, including a Land Parcel Identification System (LPIS). This spatial database is used by the decision-maker to store farmers' declarations and to check the claimed parcels (MARS, 2008; Sagris et al., 2013).

1.2 Use of remote sensing for CAP subsidy control

Control with Remote Sensing (CwRS) is a standardized control method that has been developed by the Monitoring Agricultural Resources (MARS) unit of the EC's Joint Research Centre (JRC) (Loudjani, 2013). Since 1993, the use of CwRS has been promoted as an appropriate method, equivalent to a physical farm visit, for NCPAs to check if area-based subsidies in the context of CAP are correctly granted (IES, 2013). The philosophy of CwRS is to perform OTSC of EU farms "in the office" as much as possible, given the large amount of applications that must be controlled within a short period of time and the relatively high costs of classical farm visits (Wikicap, 2012; FÖMI, 2013). CwRS relies on remote sensing data which is acquired and distributed to the NCPAs by the JRC (IES, 2013; Loudjani, 2013).

The core task of CwRS is Computer Assisted Photo-Interpretation (CAPI), whose goal is to perform both crop and area checks of a farmers' subsidy application. It is usually carried out using a time series of High Resolution (HR) imagery and at least one Very High Resolution (VHR) image (aerial orthophoto or satellite image with a pixel size < 1m) of the current control year. The former is used to check the crop occupation and the latter to verify the area of each declared parcel, since it allows to precisely identify parcel boundaries. Whenever the imagery does not allow the verification of some of the parcels listed in a farmer's application in a satisfactory way, a farm visit will have to be carried out (Wikicap, 2011a, 2013).

The CwRS program is one of the largest remote sensing programs in the world with several hundreds of satellite images acquired every year for each campaign (MARS, 2008). In 2012, it enabled control of 312 000 EU farmers' area-based subsidy applications. This represented approximately 89% of the total EU27 scheduled controls. HR images from SPOT-4, SPOT-5, IRS-P6, RapidEye, Theos and DMC were used in 2012. As for the VHR imagery, a total of 236 903 km^2 was acquired, the main sensors being GeoEye-1, WorldView-2, WorldView-1, Ikonos and QuickBird-2 (JRC, 2012). In the portuguese case, CwRS was used to perform only about 26% of all required OTSC in 2012 and 27% in 2013. SPOT satellites were the most commonly used HR image sources by IFAP in these years (Carmona, 2014).

Given the cost and time intensive nature of the CAPI process, there seems to be a current and future need for the development of cost-effective and reproducible methods to automatize crop classification. It has been pointed out that the ability to automatically classify crops using remote sensing data is essential for crop control in the context of area-based agricultural subsidies, in order to lower costs, speed up work and improve reliability compared to farm visits or even CAPI (Blaes et al., 2005; Castillejo-González et al., 2009). However, the accuracy of this kind of classification is critical and must be taken into account during the development of

such automatic classification methodologies. All declared parcels that can not be accurately classified in an automatic manner should be subjected to CAPI or farm visits (Wikicap, 2011a).

1.3 Use of remote sensing for crop identification

Multitemporal and multispectral remote sensing imagery has been widely used for crop identification in the past years (Conrad et al., 2010; Yang et al., 2011). Time series of satellite images are a cost-effective source of data to assess land cover such as agricultural crops over large areas. Time series acquired at high spatial resolution are available thanks to repetitive acquisitions by satellites such as SPOT, Landsat, IRS, and Formosat (El Hajj et al., 2009). This data allows a detailed discrimination of agricultural land use at parcel scale (Conrad et al., 2010). Several authors have recognized the benefits of this kind of data for classifying land cover, identifying crops, mapping crop rotations, and monitoring harvest and planting (El Hajj et al., 2009).

The basis for separating one crop from another is the supposition that each crop species has a specific spectral signature in a time series of multispectral images. Imagery with more spectral bands usually offers better separations for crops with similar spectral signatures. However, discrimination between crops is complex for several reasons. Agronomic factors such as variations in soil properties, fertilization, pest conditions, irrigation, planting dates, and intercropping affect and complicate crop classification. Major limitations on crop identification with satellite imagery are the similarity of plant reflectance of different crops, parcel-to-parcel variability of plant reflectance of the same crops, and the pattern of individual crop phenologies (Yang et al., 2011). In addition, to achieve a good classification result, a proper combination of spatial (pixel size several times smaller than parcel size), spectral (data on several regions in the optical spectrum) and temporal (images at several dates of annual crop development) resolution is required. Another factor that can lead to insufficient results is the conventional pixel-based classification methodology, where pixels are classified individually regardless of their spatial aggregation, i.e. where contextual information is not incorporated in the classification method. High misclassification rates of this approach are usually due to spatial variability within parcels and the presence of mixed pixels at the boundary between parcels (Peña Barragán et al., 2011; Yang et al., 2011).

To overcome these problems, object-based techniques have been increasingly used in remotely sensed image analysis. The basic idea of object-based image analysis is to use spatial objects in order to aggregate all pixels lying within an object, assigning that set of pixels to the same class in the classification process. This removes the problem of mixed pixels. There are two possible origins of the objects: objects are either "native" or are obtained by image segmentation algorithms. In the case of crop classification applications, agricultural parcels contained in the LPIS are examples of "native" objects. Object-based classifications using this kind of objects combine vector (parcel boundaries) and raster (remotely sensed images) information and can be called parcel-based classifications. Object-based image analysis can simultaneously use several

data types for classification, such as pixel values, contextual information, object features and neighborhood and hierarchical relationships (Conrad et al., 2010; Peña Barragán et al., 2011).

Several studies have been conducted on crop discrimination using object-oriented remote sensing techniques with multispectral imagery. Blaes et al. (2005) developed a strategy for crop identification in the operational context of CAP subsidy controls. This strategy relied on a parcel-based classification of multispectral and multitemporal signatures retrieved from both optical and Synthetic Aperture Radar remotely sensed imagery. They proposed the "efficiency concept" as an indicator of classification performance of the crop control system. The authors state that in an operational context, the goal of remote sensing was to identify the agricultural parcels that must be verified by farm visits. Their control system aims to maximize the "efficiency", which is defined as the number of parcels erroneously declared by a farmer in a set of suspicious parcels classified with remotely sensed imagery. Beside other results, they reached an efficiency of 24% using a hierarchical classification strategy which takes into account the variability of the spectral signatures within each crop type. The following studies rely on object-based classification, obtained by image segmentation. Castillejo-González et al. (2009) compared object-based and pixel-based classifications for identifying crops in a typical agricultural Mediterranean area, using a set of supervised classification methods and a multispectral single-date Quickbird image. Their results showed that object-based classification slightly outperformed the pixel-based methodology for the most accurate classifier (maximum likelihood) and for the non-pan-sharpened imagery, yielding an overall accuracy of 91% and 90%, respectively. The difference was substantially greater using other algorithms, e.g. the Spectral Angel Mapper classifier which resulted in an overall accuracy of 89% and 62%, respectively. Leite et al. (2011) proposed a Hidden Markov Model based technique for object-based classification of agricultural crops. To that end, they used 12 Landsat images of southeastern Brazil acquired between 2002 and 2004 and reference data provided by visual classification. In their study, the authors claimed a 93% average class accuracy in the identification of the correct crop, being, respectively, 10% and 26% superior to multi-date and single-date alternative approaches applied to the same data set. Peña Barragán et al. (2011) used vegetation indices and textural features derived from ASTER imagery to identify 13 major crops cultivated in the agricultural area of Yolo County in California, USA. The scenes were collected during three distinct growing-season periods: mid-spring, early-summer and late-summer. They combined object-based image analysis and decision tree algorithms to develop a methodology named Object-based Crop Identification and Mapping. The methodology was built in four different scenarios: combinations of three or two growing-season periods. They reported a 79% overall accuracy for the classification of the entire cropland area in the three-season period scenario. Many other studies using different combinations of sensors and classification methods can be found in literature.

1.4 Objectives of the study

The goal of this study was to develop a simple, cost-effective and reproducible method to automatize the CAPI process in an operational context with the overall objective to reduce OTSC costs and completion time. The automatization method relied on a parcel-based classification of multispectral and multitemporal land cover signatures retrieved from remotely sensed imagery, taking advantage of the object-based nature of agricultural parcels. We used 2005 validated OTSC data from IFAP and an atmospherically corrected multispectral Landsat 7 Enhanced Thematic Mapper (ETM+) time series to develop and train our methodology for the most controlled land cover classes in the portuguese Ribatejo agricultural landscape. A comparative analysis of two land cover classification methods (k-Nearest Neighbors (KNN) and Support Vector Machines (SVM)) was carried out. We took into account the importance of the accuracy of the automatic crop classification by incorporating a confidence level into our method, which can be chosen by the decision-maker. We also investigated the possibility of remotely sensed data dimensionality reduction in order to facilitate the analysis and remove possible redundancy.

2 Pattern classification theory

2.1 Overview

This chapter is designed to provide a theoretical approach to the problem of classification, starting with a brief overview of supervised and unsupervised classification and covering the Bayesian decision theory, discriminant functions, decision rules, and two classification methods, namely KNN and SVM. Accuracy assessment of classification results is also shortly covered at the end of the chapter. Readers familiar with these topics may skip this chapter and proceed directly to the chapters describing the performed study.

In general terms, a classifier assigns an object to a category based on known input variables (feature vector) describing the object. Since perfect classification is often impossible, a more general task is to determine the probability for each of the possible categories (Brenning, 2009; Duda et al., 2000). Classification can be either supervised or unsupervised. In supervised classification, the classifier is fitted to or trained on a given training dataset. This dataset consists of objects with known category membership and a set of also known variables describing the objects (Brenning, 2009). In unsupervised learning or clustering there is also training data, but the objects' category membership is unknown. The classifier forms clusters or "natural groupings" of the input variables, which is a more difficult (i.e. less well defined) problem than supervised classification (Duda et al., 2000).

Given a set of new objects with known variables, a trained classifier may be used to predict the (most likely) category membership, or some measure of the likelihood that the objects belong to a certain category (Brenning, 2009). The main goal when implementing classifiers for practical applications is an optimal performance on future unknown objects, i.e. a high generalization ability. While a highly complex classifier may allow perfect classification of the training samples, it is unlikely to perform well on new objects. This situation is known as overfitting. Choosing the right balance between maximal classifier flexibility and minimal overfitting to a limited training dataset is crucial for obtaining a good generalization ability (Belousov et al., 2002; Duda et al., 2000).

2.2 Bayesian decision theory

Bayesian decision theory is a fundamental framework to the problem of classification. This allows to quantify the trade-offs between various classification decisions using probability and the costs that are carried by such decisions. It makes the assumption that the decision problem can be described in probabilistic terms, and that all of the relevant probability values are known. In order to place our study in a formal context, we present here an introduction to the Bayesian decision theory, for more detail see Duda et al. (2000).

Let $\{\omega_1, \ldots, \omega_c\}$ be the finite set of c possible classification categories. Categories are de-

scribed probabilistically as random variables due to their unpredictable nature. Let the feature vector $\mathbf{x}$ be a d-component vector-valued random variable in a d-dimensional Euclidean space $\mathbf{R}^d$, called the feature space. For a well chosen feature space, different categories will result in different measurements of the feature vector. This variability can be expressed in probabilistic terms: the distribution of $\mathbf{x}$ depends on the true category and is expressed as $P(\mathbf{x}|\omega_j)$. This represents the category-conditional probability density function for $\mathbf{x}$, with ω_j being the true category. Let $P(\omega_j)$ be the *prior* probability that ω_j is the true category and let $P(\omega_j|\mathbf{x})$ be the *posterior* probability: the probability of the true category being ω_j given that $\mathbf{x}$ has been measured. The posterior probability can be calculated by the Bayes formula:

$$P(\omega_j|\mathbf{x}) = \frac{P(\mathbf{x}|\omega_j)\,P(\omega_j)}{P(\mathbf{x})}, \quad \forall j = 1, \ldots, c \tag{2.1}$$

In simple terms, the Bayes formula shows that by observing $\mathbf{x}$, we can convert the prior probability $P(\omega_j)$ to the posterior probability $P(\omega_j|\mathbf{x})$.

2.3 Discriminant functions and decision rules

When classifying a new object with measured $\mathbf{x}$, the probability of error by deciding category ω_j can be minimized by the following simple Bayes decision rule:

$$\text{Decide } \omega_j \text{ if } P(\omega_j|\mathbf{x}) > P(\omega_i|\mathbf{x}), \quad \forall j \neq i \tag{2.2}$$

Therefore, the result of classifying a new object with unknown category membership can be seen as a *decision*, which is the term used throughout this study. There are many different ways to represent classifiers. One of the most useful is in terms of a set of discriminant functions $g_j(\mathbf{x}), \forall j = 1, \ldots, c$. The classifier is viewed as a "machine" that computes c discriminant functions and selects the category corresponding to the largest discriminant. It assigns an object with feature vector $\mathbf{x}$ to category ω_j based on the following decision rule:

$$\text{Decide } \omega_j \text{ if } g_j(\mathbf{x}) > g_i(\mathbf{x}), \quad \forall j \neq i \tag{2.3}$$

The effect of any decision rule is to divide the feature space $\mathbf{R}^d$ into c decision regions: $\mathcal{R}_1, \ldots, \mathcal{R}_c$. If $g_j(\mathbf{x}) > g_i(\mathbf{x})$ for all $j \neq i$, then $\mathbf{x}$ is in $\mathcal{R}_j$, and the decision rule assigns $\mathbf{x}$ to ω_j. The regions are separated by decision boundaries, surfaces in feature space where ties occur among the largest discriminant functions (Duda et al., 2000).

2.4 Classification methods

In most classification applications, prior probabilities and category-conditional densities are unknown. This prevents the direct application of the presented methods, specifically the Bayes formula. There are essentially two approaches to solve this problem: parametric and nonpara-

metric estimation methods. Parametric methods assume that the forms for the probability densities $P(\mathbf{x}|\omega_j)$ are known, and use training data to estimate the unknown values of their parameters. The classic case occurs when the densities can be assumed as multivariate normal. Maximum-likelihood estimation and Bayesian estimation are two common procedures used for parameter estimation. On the other hand, nonparametric procedures can be used with arbitrary distributions and without the assumption that the form of the underlying densities are known. There are several types of nonparametric methods. One approach is to estimate the density functions $P(\mathbf{x}|\omega_j)$ from training data and use the estimated $P(\mathbf{x}|\omega_j)$ as the true density when designing the classifier. Another consists of directly estimating the posterior probabilities $P(\omega_j|\mathbf{x})$ and therefore the category membership of a given object. This approach can be exemplified by the KNN method. Furthermore, nonparametric methods exist that instead of assuming forms of probability distributions, assume forms for the discriminant functions. Training data is thus used to estimate the values of parameters of the classifier. Linear discriminant functions and the SVM technique use this type of assumption (Duda et al., 2000). Two nonparametric methods, namely KNN and SVM, are used in this study and will be further explored below.

2.4.1 k-Nearest Neighbors

The k-Nearest Neighbors (KNN) is a simple nonparametric classification method which directly estimates the posterior probability $P(\omega_j|\mathbf{x})$ for a new object from training data (Duda et al., 2000). Let $\mathcal{D}_k = \{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ be the k elements of the training data which are closest to a new object $\mathbf{x}$ in the feature space $\mathbf{R}^d$, using some distance function to measure the proximity between observations. The elements in $\mathcal{D}_k$ are called the k nearest neighbors of $\mathbf{x}$. The method simply classifies $\mathbf{x}$ by assigning it the category most frequently represented among the nearest neighbors (Carreiras et al., 2006; Duda et al., 2000). $P(\omega_j|\mathbf{x})$ is estimated by the proportion of neighbors that belong to ω_j among the nearest neighbors:

$$\widehat{P}(\omega_j|\mathbf{x}) = \frac{\#\,\omega_j}{k} \tag{2.4}$$

where $\widehat{P}(\omega_j|\mathbf{x})$ is the estimated posterior probability and $\#\,\omega_j$ is the number of neighbors with category ω_j membership. If $k = 1$, then the method is named the nearest neighbor classifier which simply assigns each object to the category of its nearest neighbor, usually resulting in overfitting the classification to the training data. Choosing a larger k can overcome this problem. The crucial parameter of the KNN classifier is therefore the number of neighbors (Carreiras et al., 2006). A large value of k leads to a more reliable estimate of $P(\omega_j|\mathbf{x})$. On the other hand, it is important to ensure that the k nearest neighbors are not overly distant from $\mathbf{x}$ (Duda et al., 2000).

The KNN method is a simple and well known classifier frequently used in various types of classification problems, effective even with large training data samples. A disadvantage of KNN is that all the objects in the training data have to be stored in memory for future classifications, since each time a new object is classified, the classifier needs to compute the distance between

that object and all other objects in the training data. For other methods, training objects are no longer required once the classifier is established. Moreover, the classifier is very sensitive to irrelevant or redundant objects since all k nearest objects contribute to the classification. It is also not clear which type of distance should be used in order to produce the best classification results (Imandoust and Bolandraftar, 2013; Marçal et al., 2005).

2.4.2 Support Vector Machines

Support Vector Machines (SVM) is a supervised non-parametric statistical machine learning technique (Mountrakis et al., 2011). The aim of the SVM algorithm is to determine a hyperplane that optimally separates two classes using training data (Kavzoglu and Colkesen, 2009). The term optimal separation refers to the goal of minimizing misclassifications (Mountrakis et al., 2011).

For the classification of d-dimensional data sets, a $(d-1)$ dimensional hyperplane is produced with SVM. In its original form with only two categories, SVM tries to locate a hyperplane that maximizes the distance from the members of each category to the optimal hyperplane. Figure 2.1 shows that there are numerous hyperplanes separating two categories. However, there is only one hyperplane that guarantees a maximum margin between the two categories, which is called the optimum hyperplane. The objects that are responsible for the width of the margin are called support vectors (Kavzoglu and Colkesen, 2009). The support vectors are the objects that are the most difficult to classify and are, at the same time, the most informative objects for the classification task. The generalization ability of the classifier is expected to increase with the margin between categories (Duda et al., 2000).

For two separable categories and following the notation introduced before, let $\{x_i, y_i\}$ $(i = 1, \ldots, n)$ be the n-sized training dataset, where $x_i \in \mathbf{R}^d$, and $y_i \in \{-1, +1\}$ is the category label (Huang et al., 2008; Kavzoglu and Colkesen, 2009). SVM searches for an optimal hyperplane defined by $w = (w_1, \ldots, w_d)$ and b such that:

$$y_i(w\, x_i + b) \geq 1, \quad \forall i = 1, \ldots, n \tag{2.5}$$

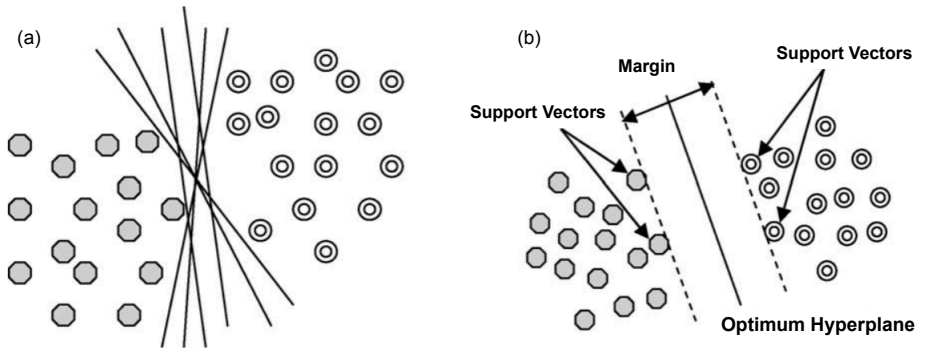


Figure 2.1: (a) Possible hyperplanes for linearly separable data. (b) Optimum hyperplane and support vectors. Source: adapted from Kavzoglu and Colkesen (2009).

If a hyperplane exists that satisfies Equation 2.5, the two categories are linearly separable. In this case, determination of the optimum hyperplane is achieved by solving the following optimization problem under the constraint in Equation 2.5:

$$\min\left(\frac{1}{2}\|w\|^2\right) \tag{2.6}$$

An assumption to the above solution is that the data points are separable in the feature space. However, in practice, data points of different categories usually overlap one another, leading to a situation with no optimal solution (Huang et al., 2002). To overcome this problem, the constraint in Equation 2.5 can be relaxed by introducing a penalty value C for misclassification errors and slack variables ξ_i are introduced in equation 2.6 as follows:

$$\min\left(\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi_i\right) \tag{2.7}$$

subject to the new relaxed constraints:

$$y_i(w\,x_i + b) \geq 1 - \xi_i \quad \text{and} \quad \xi_i \geq 0, \quad \forall i = 1, \ldots, n \tag{2.8}$$

where the penalty parameter C allows for setting a balance between the two competing criteria of margin maximization and error minimization, and the slack variables ξ_i indicate the distance of the incorrectly classified points from the optimal hyperplane (Kavzoglu and Colkesen, 2009).

The above method can be generalized to nonlinear decision functions. The SVM algorithm maps the data into a high-dimensional feature space H through a nonlinear mapping function ϕ, using a kernel function (Figure 2.2). The kernel function enables the data points to spread in a way that a linear hyperplane can be fitted in that space. Kernel functions commonly used in SVM can be aggregated into four groups: linear, polynomial, Radial Basis Function (RBF) and sigmoid kernels. RBF and polynomial kernel are the most common kernel functions used for remote sensing applications (Huang et al., 2002; Kavzoglu and Colkesen, 2009).

The SVM was described as a binary classifier, i.e. one SVM can only separate two categories. However, typical remote sensing problems usually involve separation of more than two categories.

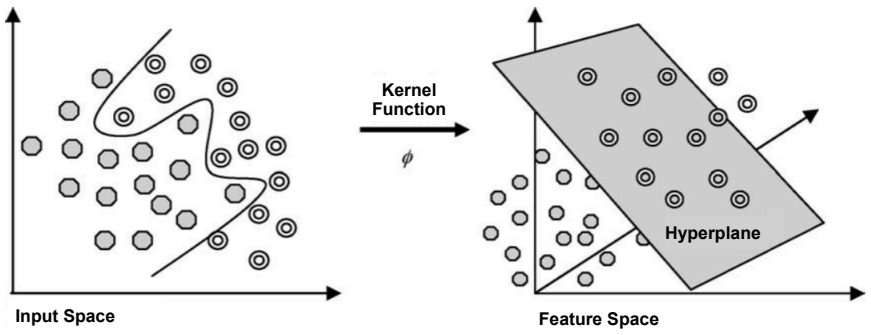


Figure 2.2: Mapping of the dataset to a high-dimensional space with a Kernel function. Source: adapted from Kavzoglu and Colkesen (2009).

Adjustments are made to the simple SVM binary classifier to operate as a multi-category classifier using strategies such as one-against-all and one-against-others (Mountrakis et al., 2011). In one-against-all strategy, a set of binary SVMs classifiers, each trained to separate one category from the rest, is applied. In one-against-one strategy, for N categories, $N(N-1)/2$ SVMs are constructed for each pair of categories. When applied to an object, each machine gives one vote to the winning category, and the object is assigned the category having most votes (Huang et al., 2002; Kavzoglu and Colkesen, 2009). Estimating posterior probabilities of SVM-based classification decisions is not a trivial task as it is, for instance, in the KNN method. Lin et al. (2007) proposed a robust algorithm which generates those estimates in a satisfactory way. An extensive tutorial on the basic ideas behind SVM can be found in Burges (1998).

SVM has been successfully used in many fields and have recently seen increased use in remote sensing applications (Huang et al., 2008; Kavzoglu and Colkesen, 2009; Mountrakis et al., 2011). One of the advantages of SVM over other machine learning algorithms such as decision trees and neural networks, demonstrated in numerous studies, is that it searches for an optimal solution to a classification problem, while decision trees and neural networks find a solution which may or may not be optimal (Huang et al., 2008). SVM are particularly attractive for remote sensing applications due to their ability to successfully handle small training datasets, often producing higher classification accuracy than more traditional methods (Mountrakis et al., 2011). Moreover, SVM-based classification has been known to be less susceptible to problems of overfitting than some other methods (Duda et al., 2000; Mountrakis et al., 2011). Even though these obvious advantages, there are also several challenges. The major difficulty regarding the applicability of SVMs is the choice of kernel functions. Some of the available kernel functions may not provide optimal SVM configuration for remote sensing applications. Studies indicate that kernels such as radial basis function and polynomial kernels applied on SVM-based classification of satellite image data produce different results. The best choice of kernel for a given problem is still a research issue (Burges, 1998; Mountrakis et al., 2011). Furthermore, the nonlinear SVM performance is affected by class imbalance, which happens when one class is represented by only a small number of training objects while other classes make up the majority (Chang and Lin, 2011).

2.5 Accuracy assessment

An important concern in remote sensing applications is to quantify the agreement between the performed classification and reference data (often called ground truth data) by performing accuracy assessment (Castillejo-González et al., 2009). In simple terms, accuracy assessment determines the quality of the information derived from remotely sensed data. This is useful for assisting in decision-making processes when remotely sensed data is used and for comparing various techniques and classification algorithms (Congalton, 1991).

Accuracy assessment is usually carried out using the error matrix approach, which consists of a square array of numbers set out in rows and columns expressing the number of objects

assigned to a particular category relative to the actual category as verified on the ground. The columns represent reference data, while the rows represent the classification decisions, with the main diagonal representing the correctly classified objects, i.e. the classification matching the reference data. Classification accuracy statistics including Overall Accuracy (OA), Producer's Accuracy (PA) and User's Accuracy (UA) can be estimated based on the error matrix. OA is the simplest statistic which indicates the percentage of correctly classified objects, computed by dividing the total correct (i.e. the sum of the main diagonal) by the total number of objects in the error matrix. An OA of 85% has been recommended as the minimum accepted accuracy. Data not reaching this level usually requires re-classification or class aggregation. PA is an estimate of the probability of a reference object being correctly classified and is calculated by dividing the total number of correct objects in a category by the column total. It is related to the commonly used omission error (error of exclusion), defined as $1 - \text{PA}$ (Congalton, 1991; Foody, 2002). If $\{\psi_1, \ldots, \psi_c\}$ is the set of c possible classification results of a classifier, than the PA can be expressed in probabilistic terms as:

$$\text{PA}_j = P(\psi_j|\omega_j), \quad \forall j = 1, \ldots, c \tag{2.9}$$

On the other hand, UA estimates the probability that a classified object actually represents that category on the ground. UA is related to commission error (error of inclusion), which is $1 - \text{UA}$ (Congalton, 1991). It is computed by dividing the total number of correct objects in a category by the row total and can be expressed probabilistically as:

$$\text{UA}_j = P(\omega_j|\psi_j), \quad \forall j = 1, \ldots, c \tag{2.10}$$

In order to obtain a robust estimate of the accuracy statistics OA, PA and UA, a technique called k-fold cross-validation can be used. This technique begins by randomly partitioning a training dataset $\mathbf{X}$ into k subsets ("folds") $\mathbf{X}_j$, $\forall j = 1, \ldots, k$, with approximately equal numbers of observations. All observations not in $\mathbf{X}_j$ are used to train a classifier, and the classifier is applied to all observations in $\mathbf{X}_j$ and assessed, resulting in an error matrix. This procedure is repeated for each holdout set $\mathbf{X}_j$, $\forall j = 1, \ldots, k$. At completion, each observation in $\mathbf{X}$ has been held-out and classified exactly once. The k error matrices resulting from each fold are then averaged and accuracy statistics OA, PA and UA are estimated based on the average matrix. This technique provides nearly unbiased estimators for the accuracy statistics if the training sample is a random sample of the population of objects (Steele, 2005).

3 Study area and datasets

3.1 Study area description

The study area is located within the portuguese province Ribatejo, northeast of Lisbon, including mainly parts of the Lisbon and the Santarém districts (Fig. 3.1). It extends over an area of about 6390 km^2 and is situated between longitudes 9° 6′ 0″ W and 8° 9′ 36″ W and latitudes 38° 43′ 47″ N and 39° 27′ 36″ N (Datum WGS84). Agriculture is the main activity in Ribatejo with a wide range of different commodities, containing some of Portugal's richest agricultural land. This agricultural area is situated within the Tagus river basin, which plays an important role for both the agricultural activity and the climate of the region. The study area is characterized by a Mediterranean climate (Csa according to the Köppen classification) with long hot and dry summers and moderate, rainy winters. On average, annual mean temperature is between 15 and 16 °C and rainfall ranges from 600 to 800 mm/year, with many crops being irrigated during the dry summer growing season (APA; IPMA, 2014).

Ribatejo can be divided in three zones: Lezíria, Bairro and Charneca. Lezíria contains the floodplain area along the Tagus river and surrounding land. Soils are very fertile and highly productive and show different textures, ranging between sand, loam and clay. Vineyards, cereals, tomato and pastures are the most common crops cultivated in this deep alluvial soil. Bairro is situated on the right margin of the Tagus, behind Lezíria (North Ribatejo). It contains poorer sandy-clayey soils with colors ranging between white, tan, red and orange. Agricultural parcels are irregularly distributed between hills and plains and the agricultural landscape is mostly shaped by vineyards, olive groves, wheat and maize. Charneca is in the southern part of Ribatejo and extends from Lezíria to the Alentejo region. Its sandy and highly heterogeneous soils are cultivated with cereals, vineyards, cork trees and rice in irrigated areas. The Tagus river is responsible for the relatively flat relief of the whole study area, with a height ranging between 0 and 200 m above sea level (APA; Carvalho, 2010; Infovini, 2014; Câmara Municipal de Santarém).

The very diverse cropping pattern found in the Ribatejo province was the main reason this area was selected for our study, along with the good availability of validated OTSC data from IFAP.

3.2 Remote sensing data

Table 3.1 shows the five optimal periods used for multitemporal image acquisition by IFAP to perform the CAPI process (Carmona, 2012). IFAP usually requests one HR image per period to JRC. These periods were chosen in order to optimally follow the annual agricultural cycle of both winter and summer crops in the typical portuguese agriculture. Beside this guideline, images for this study were selected according to image quality and the degree of atmospheric

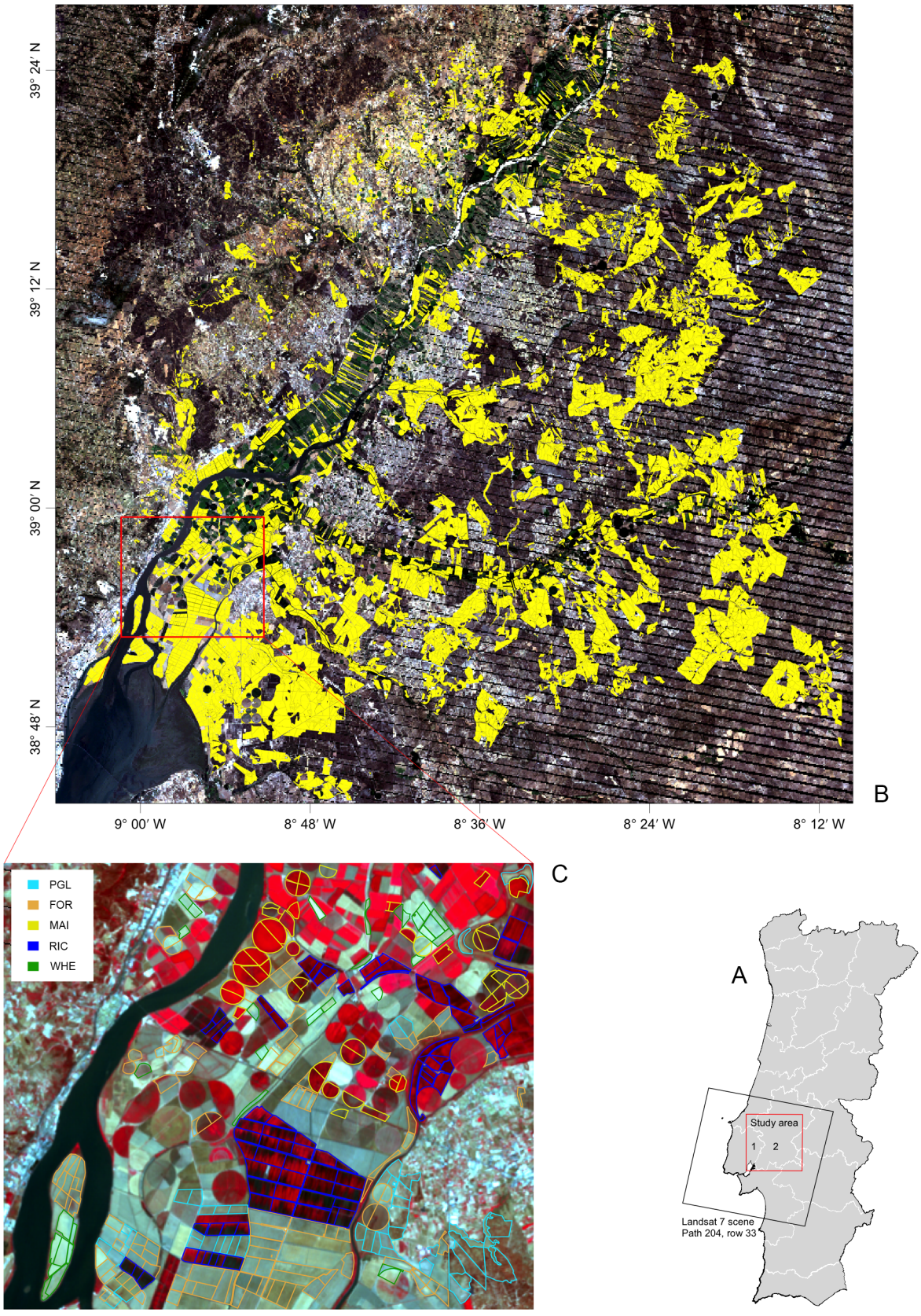


Figure 3.1: A – Study area and Landsat 7 scene location over Southern Portugal (1: Lisbon District, 2: Santarém District). B – ETM+ true color composite of the whole study area at July 8, 2005, with data loss due to the SLC-off issue being clearly noticeable. Yellow areas are the 11 852 parcels used in this work, covering approximately 1057 km^2 of the study area. C – False color composite (R = NIR, G = Red, B = Green) of a selected area in the sub-region Lezíria with parcel delineations of the five most common crop classes. Class labels can be found in Table 3.2.

Table 3.1: Optimal periods for image acquisition according to Carmona (2012) used by IFAP and acquisition dates of Landsat ETM+ images selected for this study (date formats: mm/dd and mm/dd/yyyy).

N	Optimal period	Acquisition date
Period 1	10/15 to 11/15	11/10/2004
Period 2	02/15 to 03/15	02/14/2005
Period 3	04/15 to 06/15	05/05/2005
		06/06/2005
Period 4	07/01 to 07/15	07/08/2005
Period 5	08/01 to 08/07	08/28/2005

interference, primarily due to clouds.

Based upon these criteria, six Landsat-7 ETM+ multispectral images over the study area (WRS-2 path 204, row 33) from a period between November 2004 and August 2005 were acquired through the United States Geological Survey (USGS) EarthExplorer (Table 3.1). The imagery consisted of atmospherically corrected surface reflectance data (Climate Data Records), generated from the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) tool. LEDAPS uses Moderate Resolution Imaging Spectroradiometer (MODIS) atmospheric correction routines and complex Second Simulation of a Satellite Signal in the Solar Spectrum (6S) radiative transfer models in order to generate surface reflectance (USGS, 2014b). Atmospheric correction of the image time-series is mandatory in this study since atmospheric absorption and scattering strongly affect the inter-comparability of satellite images acquired on different dates. It is highly recommended for image classification when training data from one time or place is used for classification in another time or place, a generalization which is particularly desirable in the kind of work we present here (Vicente-Serrano et al., 2008; Song et al., 2001). According to Wikicap (2011b), atmospheric effect corrections may be applied in the context of the CAP if an automatized image classification system is used in conjunction with multi-date imagery. The six visible and short-wave infrared bands (bands 1-5 and 7) with a spatial resolution of 30 m were used (see Table 5.1 for detailed information about the used Landsat-7 ETM+ bands).

The image collected for Period 2 was out of the established optimal period, but by only one day. Two good quality images were available at USGS EarthExplorer for Period 3. With roughly one month lying between these two images, we believed the classification could be improved by using the information contained in both images. The image for Period 3 should be in the first week of August, with the closest available image being at the end of that month. We believed that this image is still representative of the end of the summer crops' agricultural cycle and therefore included it in our study. The bands from all multispectral images in the temporal sequence (six bands for each image) were stacked, forming a single multi-date image with 36 bands (further called the multi-date image-stack).

Landsat data was primarily chosen due to its cost-effectiveness and good temporal coverage. In addition, Landsat ETM+ imagery was found to be appropriate for detailed large-area crop mapping, given the instruments' multiple spectral bands, which cover from visible to middle infrared wavelength regions, and its high 30 m spatial resolution (Wardlow et al., 2007). The possibility to acquire pre-processed images, with atmospheric correction and converted to surface

reflectance, further improves the usability of Landsat data. Another good reason for using Landsat data is that Landsat 8 imagery with the same LEDAPS pre-processing is expected to become available in 2014 (USGS, 2014a). This ensures a continuous high-quality availability of remotely sensed data to be used in the near future.

All the acquired images were affected by the so-called Scan Line Corrector (SLC)-off issue, which results in the loss of approximately 22% of the normal scene area for Landsat ETM+ acquisitions after May 31, 2003. Data loss varies in width from one pixel or less near the center of the image to 14 pixels along the eastern and western edges of the image. However, the SLC-off issue has no impact on the quality of the remaining valid pixels (Chander et al., 2009). Our study area covered only a relatively small area of the scenes, including their central part, and was therefore not highly affected. A "fill-value" was assigned to all pixels with missing data which were ignored in further processing steps.

3.3 Agricultural parcels and land cover classes

Data from agricultural parcels controlled in 2005 by IFAP in the context of CAP was used in this study to develop and train our CAPI automatization method. The following relevant information about each parcel was provided: spatial delineation, area and validated land cover. The year 2005 was chosen due to the relatively high control rate in Ribatejo in that year, allowing us to analyze the high crop diversity mentioned before. Being controlled in 2005 means that the data contains both winter 2004/2005 crops and summer 2005 crops, which is the reason the Landsat image acquired in November 2004 was needed.

A total of 32 062 agricultural parcels lying within the study area were initially extracted from the LPIS and provided for this study by IFAP. For each parcel, an extraction of all the multi-date image-stack pixels lying within the parcel's boundary was performed, resulting in a data structure where all parcels contained multitemporal and multispectral information. Only pixels lying entirely within a parcel were considered in order to avoid mixed pixels at the boundary that could affect the classification performance.

A subset of the initial parcels was then selected for further analysis based on two criteria. The first and obvious criterion is that all the parcels had to contain at least one entire pixel. Secondly, only parcels containing land cover classes accounting for approximately 95% of the total parcel area were included in the subset. In detail, 12 classes (Table 3.2) covered 94.4% of the total parcel area of 111 963 ha, with the remaining 5.6% being covered by 48 classes. This resulted in a total of 11 852 parcels used for this work, representing 105 702 ha of the study area. We preferred the global designation "land cover class" over "crop" due to the fact that some classes, in particular Fallow and Non used area, can not be seen as crops. The designation "crop class" can be used to refer to one of the remaining 10 classes. Non used area is a class with agricultural inactivity, differing from Fallow which, unlike Non used area, is meant to produce agricultural goods in the future. However, Non used area can be declared by a farmer and had therefore to be included in this work. Note that, strictly speaking, the used 11 852 parcels do

Table 3.2: Summary of the used land cover classes. More than half of the area is covered by Permanent grassland. Note the great variability in parcel size found in the dataset (measured in standard deviation of parcel area).

Land cover class	Class label	Number of parcels	Total area (ha)	Relative area (%)	Average area (ha)
Permanent grassland	PGL	4051	59243	56.0	14.6 ± 13.9
Forage crops	FOR	1708	12183	11.5	7.1 ± 8.2
Maize	MAI	1190	8916	8.4	7.5 ± 6.5
Rice	RIC	799	6501	6.2	8.1 ± 6.9
Fallow	FAL	1473	5944	5.6	4 ± 5.1
Wheat	WHE	578	2710	2.6	4.7 ± 4.8
Poor grassland	POG	330	2707	2.6	8.2 ± 10.4
Vineyard	VYA	578	2197	2.1	3.8 ± 3.6
Non used area	NUA	329	1727	1.6	5.2 ± 9.8
Barley	BAR	275	1645	1.6	6 ± 6.4
Oat	OAT	282	1189	1.1	4.2 ± 5.4
Olive grove	OLI	259	741	0.7	2.9 ± 3

not reflect with certainty the most cultivated crops in the region, but the most controlled crops in 2005 in the study area. According to the notation introduced in Section 2.2 and followed throughout this study, $\{\omega_1, \ldots, \omega_{12}\}$ are the $c = 12$ possible land cover class decisions.

4 Methods

4.1 Overview

The CAPI automatization method is a classification problem which can be handled using the theory explored in Chapter 2. In simple terms, the goal of this classification is to take a reliable decision regarding the land cover class of a future unknown agricultural parcel. The major differences of this method compared with standard pixel-based image classification are: 1) we included a confidence level defined a priori by the decision-maker in order to choose the level of reliability of the automatization system, and 2) we took advantage of the object-based nature of the agricultural parcels in farmers' declarations to carry out a parcel-based classification. These features are intimately related to the nature of the application.

The method is divided in two subsequent phases: 1) development and calibration phase and 2) application phase. Figure 4.1 shows an overview of the processing steps of the development and calibration phase. In this phase, a classification training dataset is first derived from remotely sensed data, consisting of agricultural parcels with known land cover class membership as detailed in Section 4.2. A feature space dimensionality reduction method is also presented in this section. Section 4.3 and 4.4 describe the training and accuracy assessment of two classifiers using a 10-fold cross-validation technique for optimal classifier selection. In Section 4.5 we propose our method for taking classification decisions of future unknown parcels in function of the desired reliability of the classification system. The application phase corresponds to the practical application of the method, represented in Figure 4.3 and further detailed in Section 4.5.

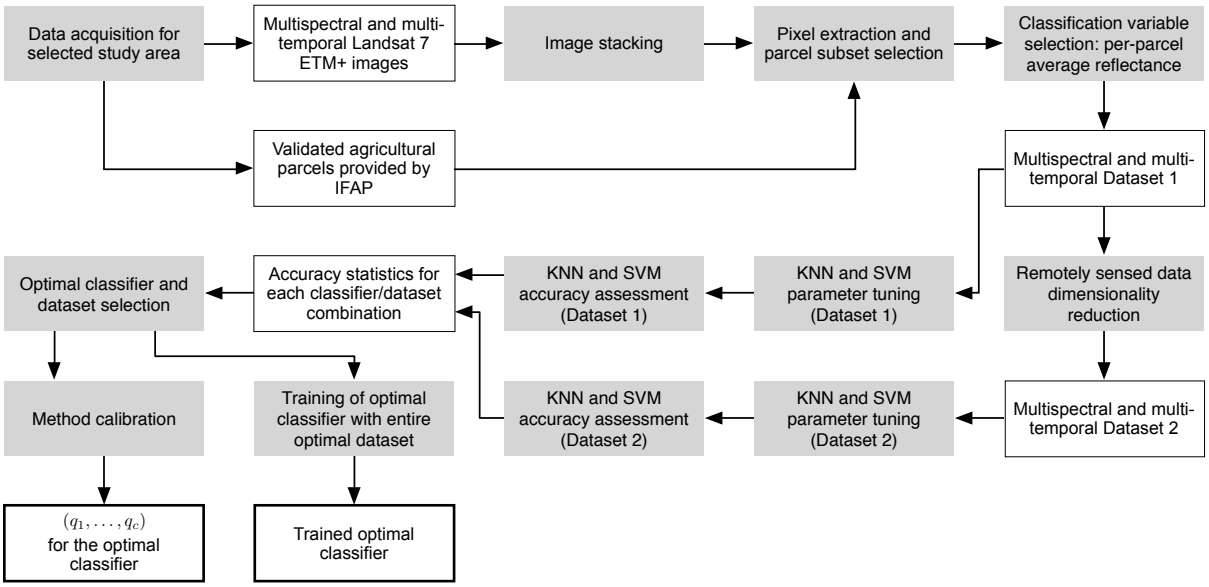


Figure 4.1: Overall workflow of the CAPI automatization method development and calibration phase. Grey backgrounds represent processes while white backgrounds stand for datasets/results. Bold borders indicate outputs to the application phase (see Figure 4.3).

All steps were performed using the freely available software environment R, version 3.0.2 of the 64-bit version, in conjunction with the graphical user interface RStudio, version 0.98.945 (R Development Core Team, 2014; RStudio, 2014). Several libraries (so-called packages) that extend the base functionality of the software were used to implement the method, such as `maptools` (Bivand, 2014), `sp` (Pebesma et al., 2013), `raster` (Hijmans, 2014), `rgeos` (Bivand et al., 2014b), `rgdal` (Bivand et al., 2014a) and `proj4` (Urbanek, 2014) for spatial analysis and mapping of vector and raster data. Additional libraries are mentioned in the respective sections. R software has recently gained increased popularity for spatial modeling and geocomputing, including remote sensing applications, due to integrations of geocomputing capabilities in form of freely available packages (Brenning, 2009).

4.2 Feature space and dimensionality reduction

The pre-processing of the acquired datasets, namely remote sensing data and validated agricultural parcels, resulted in a data structure where all parcels contained multitemporal and multispectral information. This dataset served to derive training and reference data in form of classification variables by spatially aggregating spectral information, in order to allow a unique classification for each parcel (parcel-based classification). Per-parcel average surface reflectances (the average of all pixels found within a parcel) were used as classification variables, equivalent to a multitemporal spectral signature per object. Variables were computed for the 11 852 parcels, resulting in a dataset $\mathbf{X}$ with 36 predictor variables (6 images with 6 bands each), denominated Dataset 1. It is important to highlight the fact that each variable consists of a combination between spectral bands and image acquisition dates. Pixels with "fill-values", i.e. pixels with no information due to the SLC-off issue, were ignored in average calculations. Further classification variables, specifically per-parcel standard deviation of reflectance and average NDVI, were also computed initially but discarded for not improving the overall classification accuracy. Adopting the conventions introduced in Section 2.2, for each parcel, $\mathbf{x}$ is a vector containing the 36 classification variables in the feature space $\mathbf{R}^{36}$ which consists of a combination of spectral bands and dates. The posterior probability $P(\omega_j|\mathbf{x})$ is therefore the probability that ω_j is the true class of a parcel given a measured multitemporal spectral signature, for all $j = 1, \ldots, 12$ classes.

As a next step, dimensionality of the feature space $\mathbf{R}^{36}$ was reduced by selecting a subset of the original classification variables. Dimensionality reduction is a common pre-processing step in supervised classification of hyperspectral images for various applications such as land cover mapping and can also be used with multispectral data (Damodaran and Nidamanuri, 2014). Several reasons led us to perform dimensionality reduction: 1) reduction of possible redundancy in Dataset 1 (e.g. due to linear dependencies), 2) excluding variables could lead to a reduced number of required multispectral images for classification, thus facilitating the practical application of the method in an operational context, 3) overcoming of the potential problem of overfitting, and 4) diminution of required computational effort in both calibration

and application phase of the method, given the large training sample size of 11 852 parcels. The idea was to choose a subset denominated DATASET 2 of the whole DATASET 1 with a relatively small loss of the information stored in the original variables. We chose a very simple method which conserves the original spectral information of the variables, based on Principal Component Analysis theory.

The method is processed in two steps. Step 1 consists in successively deleting variables which contribute the most to the occurrence of linear dependencies within the original dataset. Let p be the dimension of the original dataset. The method begins by computing the correlation matrix $\mathbf{S}_{p\times p}$ of the original data. Eigenvalues $\boldsymbol{\lambda}$ of $\mathbf{S}$ are than calculated. Note that $\mathbf{S}$ is normal since it is symmetric. The condition number of $\mathbf{S}$ is computed as follows for normal matrices:

$$k(\mathbf{S}) = \left| \frac{\lambda_{\max}(\mathbf{S})}{\lambda_{\min}(\mathbf{S})} \right| \tag{4.1}$$

where $\lambda_{\max}(\mathbf{S})$ and $\lambda_{\min}(\mathbf{S})$ are maximal and minimal eigenvalues of $\mathbf{S}$, respectively. The pseudocode in Algorithm 4.1 summarizes the processing in Step 1. The least important variable which is deleted from $\mathbf{X}$ is the variable $\mathbf{x}_j$ with the highest coefficient in absolute value in the eigenvector associated to the smallest eigenvalue of $\mathbf{S}$, according to Jolliffe (2002, pp. 138). This step is implemented in a slightly different form in the `subselect` package available for the software environment R (Cadima et al., 2013).

In Step 2, the relationship between the number of deleted variables and respective $k(\mathbf{S})$ is analyzed in order to choose the number d of variables in DATASET 2. This selection is made based on a trade-off between two competing criteria: 1) the condition number should be as low as possible since that means that the problem is well-conditioned, and 2) a reasonable number of variables must be selected, i.e. only variables which are redundant should be removed. Therefore, d was fixed at a point where $k(\mathbf{S})$ started to approach the minimum value of 1. The d selected variables for DATASET 2 are equivalent the last d deleted variables in Step 1, since these are considered to best represent the variability in the original dataset.

Note that this methodology does neither take the nature of the variables into account, nor the information about the membership of any observation to a particular class. Using only the correlations between variables, it may thus not be able to choose the most relevant band/date combinations from a remote sensing point of view. In order to evaluate the impact of

Algorithm 4.1 Step 1 of the dimensionality reduction method.

Input: Dataset with variables $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$
Output: Condition numbers $k_1 > \cdots > k_p$
  for $i = 1$ to p **do**
    Compute $\mathbf{S} = \text{cor}(\mathbf{X})$
    $\lambda \leftarrow$ eigenvalues of $\mathbf{S}$
    $k_i \leftarrow |\max(\lambda)/\min(\lambda)|$
    Choose the least important variable $\mathbf{x}_j$
    Remove $\mathbf{x}_j$ from $\mathbf{X}$
  end for

dimensionality reduction in classification performance, both DATASET 1 and DATASET 2 were used for classifier training and respective accuracy assessment.

4.3 Classification parameter tuning

Parcel-based land cover classification was carried out using KNN and SVM classifiers with DATASET 1 and DATASET 2 as training data. Training and classification, as well as classifier accuracy assessment, was carried out using a 10-fold cross-validation, as discussed in the next section. In the classification process, each object is assigned a class membership and the corresponding posterior probability value is returned. The usage of classification methods requires the setting of classifier-specific parameters, which optimal values are unknown. Therefore, an optimum parameter search must be performed for both the KNN and the SVM classifier, known as tuning of the parameters.

For the KNN method, the number of neighbors k for nonparametric classification is the crucial parameter. In this study, the optimal k was tuned by maximizing the overall test set accuracy based on 10-fold cross-validation. We ran the tuning algorithm for $k = (1, \ldots, 20)$ using both DATASET 1 and DATASET 2 as input. For each of the training datasets, the k which maximized the respective overall accuracy was chosen, k_1 and k_2, respectively. The Euclidean distance was used for distance measurement. The KNN method and respective tuning algorithm are implemented in the `kknn` package for the R software (Schliep and Hechenbichler, 2014) .

As for SVM, we used the `e1071` package (Meyer et al., 2014) which offers an interface to the LIBSVM program developed by Chang and Lin (2011). For multi-category classification, this library uses the one-against-one approach. The LIBSVM program has several parameters that need to be defined, being the selection of a kernel function the first step. The performance of SVMs varies depending on the choice of the kernel function and its parameters (Burges, 1998). The RBF kernel function was used because it has fewer parameter values to define and has been found at least as robust as other kernel types for remote sensing applications (Huang et al., 2008). It is defined as follows:

$$K(x_i, x_j) = e^{-\gamma(x_i - x_j)^2} \tag{4.2}$$

where γ is the kernel width. The SVM penalty parameter C and γ need to be defined for the usage of the classifier with the selected kernel function. Since it is not clear which pairs of parameters produce the best classification result, the values for C and γ were systematically changed in order to perform optimum parameter search. After some initial tests, we ran a parameter tuning algorithm for all combinations between $C = (0.5, 1, 1.5, 2, 2.5, 3)$ and $\gamma = (0.4, 0.6, 0.8, 1, 1.2, 1.4)$. For each combination, the algorithm returned the overall test set accuracy of the SVM model through 10-fold cross-validation. This procedure was applied to DATASET 1 and DATASET 2. As done with the KNN classifier, for each of the training datasets, the combination which maximized the respective overall accuracy was chosen for classification, (C_1, γ_1) and (C_2, γ_2), respectively.

4.4 Optimal dataset and classifier selection

We quantified the agreement between parcel classification and true validated parcel land cover. Accuracy assessment was performed to identify the optimal combination of classifier (KNN and SVM) and dataset (DATASET 1 and DATASET 2) to be used for the CAPI automatization method. This assessment also allowed for evaluating whether the applied dimensionality reduction can be successfully used or not.

Accuracy assessment was undertaken using the error matrix approach described in Section 2.5. Robust estimates of Overall Accuracy (OA), Producer's Accuracy (PA) and User's Accuracy (UA) for each combination of classifier and dataset were obtained using 10-fold cross-validation ($k = 10$). Beside the mentioned accuracy statistics, this technique also provided a dataset labeled OUTPUT, containing the true and the estimated land cover class memberships for all 11 852 parcels, as well as the associated posterior probabilities for usage in the next step.

4.5 Method calibration, assessment and application

In this section we present the developed CAPI automatization method and show its application in an operational context. The optimal classifier/dataset combination found in the accuracy assessment step was used for calibration and assessment of the method. As outlined before, it is crucial to take into account the accuracy of automatic land cover classification in the operational context of CAP subsidy controls, in order to build a method that produces reliable results. However, for such land cover classification system, assessment measures including OA, PA and UA are not of interest. Given the objective of a subsidy control system, this type of system requires a binary classification assessment measure: to either accept a classifier's decision or to reject it. The degree of the method's reliability should be in some way measurable and adjustable through a "confidence level" that we describe below.

Some classification results are more reliable than others, as revealed by their respective posterior probabilities $P(\omega_j|\mathbf{x})$. The question that arises is the following: which value of probability can be considered large enough to ensure that the classification is reliable? Our method attempts to answer that question by defining a minimum posterior probability for each land cover class, $\{q_1, \ldots, q_c\}$, above which a classification can be considered reliable. In short, it is not only necessary to classify an agricultural parcel with unknown land cover, but also to accept or reject that decision with a chosen confidence level. Accepting a classification decision means that the corresponding parcel is classified automatically. The proposed method establishes the simple following decision rule:

$$\text{If } \omega_j \text{ is the decision, then } \begin{cases} \text{Accept decision} & \text{if } P(\omega_j|\mathbf{x}) \geq q_j \\ \text{Reject decision} & \text{if } P(\omega_j|\mathbf{x}) < q_j \end{cases} \tag{4.3}$$

where $P(\omega_j|\mathbf{x})$ is the posterior probability of the classifier and q_j is the minimum posterior

probability for land cover class j. A low q_j implies that all classification decisions ω_j will be accepted, while $q_j = 1$ implies that only decisions ω_j with posterior probability of 1 will be accepted by the classification system. In general terms, a high q_j indicates that it is difficult to take a confident decision regarding land cover class j.

Given the decision rule in Equation 4.3, the crucial task of the method is estimating $\mathbf{q} = (q_1, \ldots, q_c)$ that guarantees the desired reliability of the system, which we call the calibration of the method. The rationale behind our approach is that the attempt to classify *all* parcels will inevitably lead to a poor classification accuracy, at least for some classes. Therefore, the goal is to select a sub-population of parcels $\mathbf{R}_q$ for which a high accuracy is achievable. We call confidence level λ to the minimum $\mathrm{UA}_j = P(\omega_j|\psi_j)$ over all classes j, which can be described as an "overall UA". Given a confidence level λ chosen by the decision-maker, the task is therefore to select the largest sub-population $\mathbf{R}_q$ for which $\mathrm{UA} \geq \lambda$ for all classes. The criterion to include or exclude a given parcel from $\mathbf{R}_q$ is that the corresponding feature vector $\mathbf{x}$ satisfies the condition $P(\omega_j|\mathbf{x}) \geq q_j$. In order to illustrate this approach, Figure 4.2 shows a schematic representation of $\mathbf{R}_q^1$ and $\mathbf{R}_q^2$ for two arbitrary classes with corresponding arbitrary q_1 and q_2. According to the decision rule in Equation 4.3, the accepted classification decisions ω_1 are $D \cup G$ and rejected classification decisions ω_1 are $C \cup F$. As for class 2, accepted decisions are $A \cup B$ and rejected decisions are E only.

In general, $\mathbf{R}_q^j = \{\mathbf{x} : P(\omega_j|\mathbf{x}) \geq q_j\}$ and $\mathbf{R}_q = \cup_j \mathbf{R}_q^j$. Given a classification confidence level $\lambda \in [0, 1]$, the method estimates the optimum $\mathbf{q} = (q_1, \ldots, q_c)$ such that in $\mathbf{R}_q$:

$$\mathrm{UA}_j \geq \lambda, \ \forall j \tag{4.4}$$

The pseudocode in Algorithm 4.2 summarizes the estimation of $\mathbf{q}$ in practical terms. It is important to emphasize the usage of the UA for method formulation rather than the PA. UA is the probability that a classified parcel was actually correctly classified. This corresponds

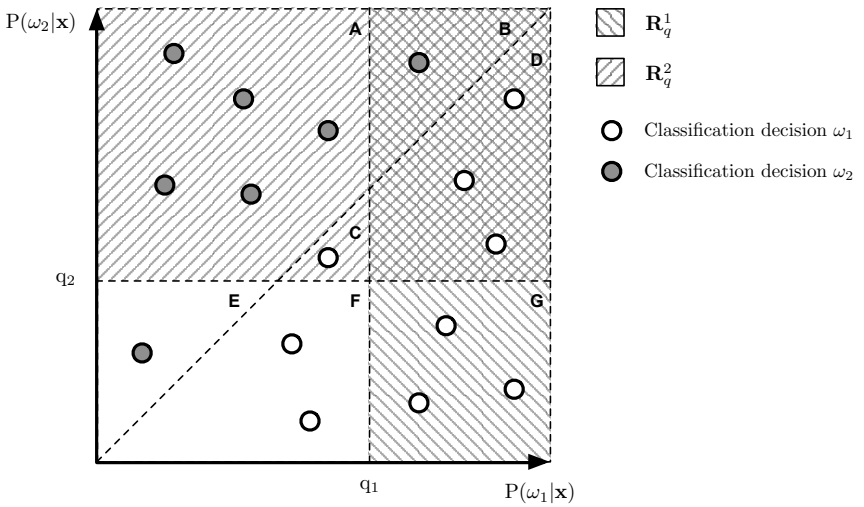


Figure 4.2: Schematic representation of $\mathbf{R}_q^1$ and $\mathbf{R}_q^2$ for two arbitrary classes and arbitrary objects. Each letter stands for the set of objects lying within the corresponding region. Note that in this example we consider that the overall number of classes is larger that 2, since otherwise $P(\omega_1|\mathbf{x}) + P(\omega_2|\mathbf{x}) = 1$.

Algorithm 4.2 Method calibration.

Input: Dataset OUTPUT (Section 4.4) and confidence level λ
Output: Probabilities $(q_1, \ldots, q_c)$
  for $j = 1$ to c **do**
    $S_j \leftarrow$ subset of parcels with decision ω_j from OUTPUT
    $UA_j \leftarrow$ proportion of S_j correctly classified
    while $UA_j < \lambda$ and S_j is not empty **do**
      remove parcel with the lowest $P(\omega_j|\mathbf{x})$ from S_j
      recompute UA_j
    end while
    if S_j is not empty **then**
      $q_j \leftarrow$ minimum $P(\omega_j|\mathbf{x})$ among all parcels in S_j
    else
      q_j is undefined
    end if
  end for

precisely to the goal of the control system which aims at maximizing the number of correctly classified parcels. Therefore, UA is the probability which needs to be maximized in order to guarantee that CAP funds are properly spent. In the proposed method, this is achieved by setting a high confidence level. On the other hand, PA is the probability that a reference parcel was correctly classified, which is not of interest for a control system. A low value of PA only means that the corresponding land cover class is difficult to separate from the remaining classes.

For the purpose of assessing our method, we fixed a confidence level and estimated the accuracy statistics OA, PA and UA only for the parcels that could be classified automatically by the method with the chosen λ. Moreover, we calculated the proportion of automatically classified parcels, with the goal of providing an overview of the practical applicability of the method in each class. Let $U_j = \{\text{Automatically classified parcels in class } j\}$ and $V_j = \{\text{Parcels with decision } \omega_j \text{ either accepted or rejected}\}$. In terms of the representation in Figure 4.2, $U_1 = D \cup G$ and $V_1 = D \cup G \cup C \cup F$. We define the Automatic Classification Proportion (ACP) for each class as follows:

$$\text{ACP}_j = \frac{|U_j|}{|V_j|} \tag{4.5}$$

where $|U_j|$ is the dimension of set U_j and $|V_j|$ is the dimension of set V_j. The overall ACP is computed by simply dividing the total number of automatically classified parcels by the total number of parcels. Note that the ACP does not provide any information about the accuracy of the classification, it only measures the relative amount of classified and not-classified parcels.

The method is designed to be used by a control agency in an operational context in the so-called application phase once the optimal classifier and the method were successfully trained and calibrated, respectively. This phase can be outlined by the general overview in Figure 4.3, including the required inputs and possible outputs. In this work, the method was both calibrated and assessed using data from 2005. In an operational context, the intent is to use input data

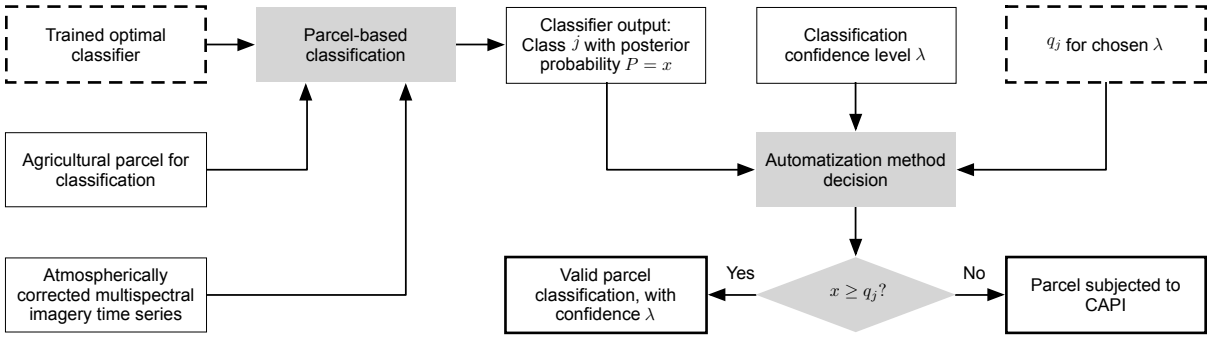


Figure 4.3: Application phase of the CAPI automatization method for a single parcel classification. Grey backgrounds represent processes while white backgrounds stand for inputs/results. Dashed borders stand for inputs from the development and calibration phase. Bold borders indicate the two possible final results.

from different years for calibration and application. For instance, automatic classification can be applied in a given year by calibrating it with data from the previous year, namely remotely sensed imagery and control data. In the application phase, beside fixing the required confidence level, the control agency is in charge of providing delineations of the parcels to be classified and an atmospherically corrected multispectral imagery time series for the control year as inputs. As mentioned before, any agricultural parcel that is rejected by the method for not meeting the desired reliability has to be subjected to CAPI. The remaining parcels can be controlled automatically using the proposed CAPI automatization method.

5 Results and discussion

5.1 Dimensionality reduction

The dimensionality reduction method produced the relationship between number of deleted variables and respective correlation matrix condition number plotted in Figure 5.1. As expected, the correlation matrix of DATASET 1 was highly ill-conditioned, as shown by the high $k(\mathbf{S})$ rounding 10^4. This effect is caused by linear dependencies between variables which represent very similar spectral information. An example of this can be two variables from the same band and from two images taken within a short period of time, such as June 6 and July 8. If the evolution of the crops' phenological cycle is not pronounced in this period, then these variables will most likely contain redundant information. Successive deletions of variables according to the described method reduced the problem of ill-condition. The condition number decreased strongly with the first deletions and approached the minimum value 1 as more and more variables were removed from the original dataset. It started to stabilize at around 25 deleted variables, taking the value of approximately 83. Based on the shown relationship, we assumed that selecting a third of the original 36 variables satisfied both the criteria of selecting a reasonable number of variables and having a relatively low condition number. $k(\mathbf{S})$ with 12 variables is approximately 53, only 0.6% of the initial value of 9307.

Table 5.1 shows the 12 variables selected by the method in terms of ETM+ band and acquisition date combinations. Analyzing the selected spectral bands, the results make sense from a remote sensing perspective, given that all variables representing the Near Infrared (NIR) spectral region (Band 4) were selected and that the red spectral region (Band 3) was the second most contributing band. These two bands form the so-called red edge which is the rapid change in reflection between red and NIR wavelengths. The origin of the red edge is the fact that vegetation absorbs most of the incident sunlight in the red spectral region and reflects large proportions of the sunlight in the NIR region. It is therefore widely used for identification and discrimination of photosynthesizing plants via remote sensing (Usha and Singh, 2013). Regarding image acquisition dates, the results of the performed dimensionality reduction show that all dates contributed equally to represent the variability of the original dataset, with two variables being selected in

Table 5.1: Landsat ETM+ band/date combinations for DATASET 2 selected by the dimensionality reduction method. Two variables per acquisition date and all variables representing the NIR spectral region (Band 4) were selected. Detailed information about each Landsat ETM+ band is provided.

Landsat band (spectral range)	Nov 10	Feb 14	May 5	Jun 6	Jul 8	Aug 25
Band 1 (0.452 – 0.514 µm: Blue)						X
Band 2 (0.519 – 0.601 µm: Green)						
Band 3 (0.631 – 0.692 µm: Red)	X		X	X		
Band 4 (0.772 – 0.898 µm: NIR)	X	X	X	X	X	X
Band 5 (1.547 – 1.748 µm: SWIR)					X	
Band 7 (2.065 – 2.346 µm: SWIR)		X				

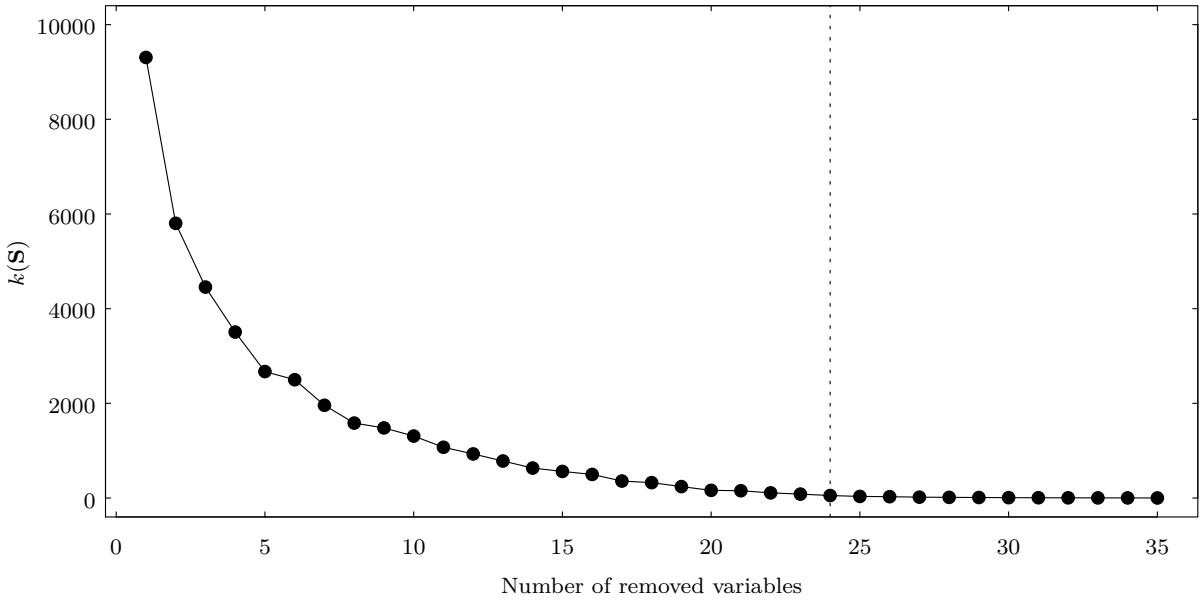


Figure 5.1: Relationship between the number of removed variables by the dimensionality reduction method and the condition number of correlation matrix **S**, necessary for selection of Dataset 2 dimensionality. The initial matrix **S** was highly ill-conditioned. 12 classification variables were maintained, thus removing 24 variables, as indicated by the dashed line.

each date. Multispectral images from all acquisition dates are therefore necessary, eliminating the possibility of reducing the amount of images to acquire for land cover classification.

The overall variable selection is fairly balanced in terms of bands and dates. This is a promising outcome for the next step which consists of evaluating whether Dataset 2, containing the 12 selected variables, is suitable for performing land cover classification.

5.2 Optimal classifier and dataset selection

Classification and accuracy assessment was carried out for all 4 combinations of classifiers and datasets to compare the respective classification performances. The main goal of this process was choosing the optimal classifier and dataset for the automatization method. Parameter tuning was performed as a first step, resulting for KNN in $k = 7$ neighbors for both datasets. An odd number of neighbors has the advantage of avoiding ties in the classification process. As for SVM with Dataset 1, penalty parameter C was 1.5 and kernel width γ was 0.4. For Dataset 2, C was 2.5 and γ remained at 0.4. This indicates that with reduced dimensionality, we need to increase the penalty for misclassification errors in order to achieve good classification results.

Table 5.2 summarizes the classification accuracy assessment results in terms of OA and detailed UA and PA for each land cover class, estimated by 10-fold cross-validation. The results reveal that based on OA, the SVM classifier outperforms KNN with both datasets, with the best result of 68.35% being yielded when SVM classifier is used with Dataset 1. Analyzing the effect of dimensionality reduction, we found that classification using the original Dataset 1 resulted in OA values of 66.26% and 68.35% for KNN and SVM, respectively. These results are slightly superior to results found with Dataset 2, 64.74% and 68.11% for KNN and SVM,

Table 5.2: Summary of the average accuracy statistics UA, PA and OA estimated by 10-fold cross-validation for all combinations of classifiers and datasets. Based on OA, the SVM classifier outperforms KNN with both datasets. SVM-2 was chosen as the optimal classifier/dataset combination for usage in the automatization method. 1 stands for DATASET 1 and 2 stands for DATASET 2.

Class	KNN–1		KNN–2		SVM–1		SVM–2	
	User	Prod	User	Prod	User	Prod	User	Prod
Permanent grassland	0.64	0.92	0.62	0.91	0.72	0.87	0.68	0.89
Forage crops	0.47	0.34	0.44	0.33	0.45	0.47	0.48	0.41
Maize	0.93	0.95	0.93	0.95	0.83	0.97	0.90	0.96
Rice	0.99	0.97	0.98	0.97	0.99	0.94	0.98	0.97
Fallow	0.50	0.30	0.49	0.28	0.50	0.43	0.49	0.42
Wheat	0.69	0.72	0.70	0.66	0.76	0.72	0.76	0.74
Poor grassland	0.23	0.06	0.29	0.05	0.32	0.06	0.30	0.04
Vineyard	0.67	0.75	0.66	0.72	0.78	0.73	0.75	0.74
Non used area	0.51	0.21	0.46	0.19	0.53	0.24	0.45	0.18
Barley	0.70	0.51	0.68	0.51	0.74	0.60	0.72	0.61
Oat	0.44	0.27	0.39	0.21	0.55	0.32	0.53	0.26
Olive grove	0.42	0.12	0.39	0.10	0.49	0.20	0.60	0.20
Overall accuracy (%)	66.26		64.74		68.35		68.11	

respectively. The SVM classifier deals particularly well with the reduced dataset, as shown by the small difference between the two SVM results ($\Delta = 0.24\%$). With SVM, some classes reveal higher classification accuracies when classified with DATASET 2: Permanent grassland, Rice, Wheat, Vineyard and Barley show higher values of PA while Forage crops, Maize and Olive grove reveal higher values of UA. Therefore, we considered that the dimensionality reduction was successful. Given these results and the advantages of using the reduced dataset, we chose the SVM classifier with DATASET 2 as the optimal combination. In addition, the SVM method has the advantage of being less susceptible to problems of overfitting, which is crucial for the required generalization ability of the method.

In general terms, OA results were rather unsatisfactory for all classifications, clearly below the recommended minimum value of 85% (Foody, 2002). Furthermore, some land cover classes yielded very low accuracies which makes classification of those classes impossible without further improvements of the process. This confirms that attempting to classify all parcels will lead to poor results. This is the reason we developed the CAPI automatization method which only automatically classifies a subset of parcels that guarantees a desired overall UA.

Figure 5.2 shows an average per-parcel multitemporal spectral signature for each land cover class, build with the variables from the dimensionality reduced DATASET 2. Spectral signatures are useful to visually understand the reason behind the fact that some classes are more difficult to classify than others. Two crops with similar multitemporal spectral signatures are characterized by both similar interaction with the incident sunlight and development cycle throughout a given year. Land cover classes with similar spectral signatures are difficult to separate by a classifier, which reflects in poor classification results of those classes. An analysis of the signatures reveals that some classes, such as Permanent grassland, Poor grassland and Non used area show very similar behavior. In fact, we can see in Table 5.2 that those classes show low values of UA and/or PA. Barley and Wheat are also similar, which makes sense since both are winter cereal

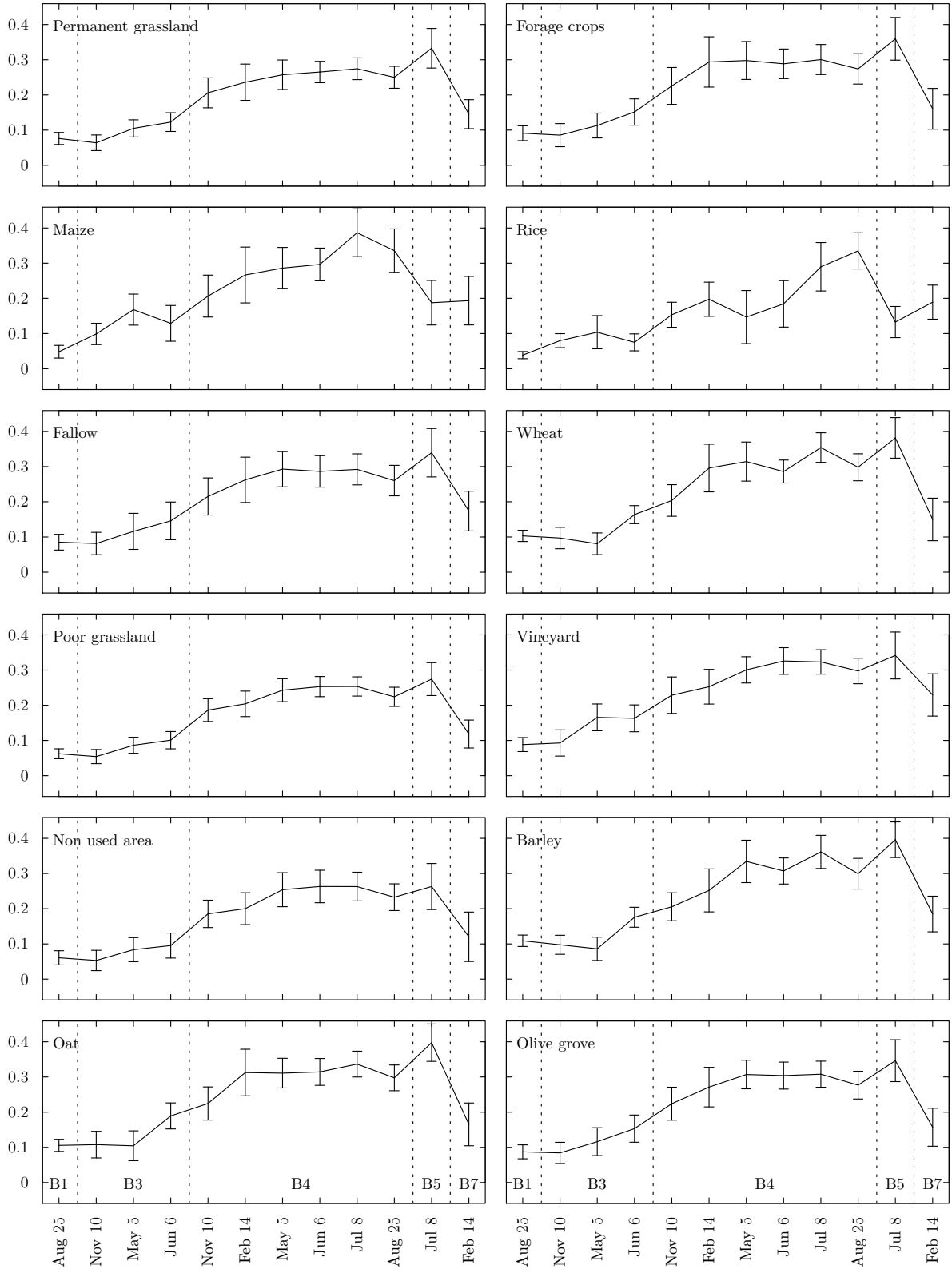


Figure 5.2: Average per-parcel multitemporal spectral signature for all 12 land cover classes. B1 – Landsat ETM+ Band1, B3 – Band 3, B4 – Band 4, B5 – Band 5, B7 – Band 7. Error bars are standard deviations. Y-axis shows parcel reflectance and x-axis shows acquisition dates. Some classes such as Permanent grassland, Poor grassland and Non used area reveal similar signatures and are therefore difficult to classify by any classifier. Other classes such as Maize, Rice and Vineyard show different behaviors and are thus expected to be easily classified.

crops. Other classes, for instance Maize, Rice and Vineyard, show quite unique signatures, which enables a good classification performance of those crops. This discussion will be further detailed in the next section.

5.3 Automatic land cover classification

5.3.1 Confidence level selection

We start with an overview of the results by analyzing the effect of the confidence level on the proportion of agriculture parcels that can be classified automatically. This information is crucial to take a decision regarding the confidence level do adopt for automatic classification. The method was calibrated which each λ between 50% and 100%, with a step of 5%, using the classification results from the chosen SVM classifier with Dataset 2. ACP and overall ACP were estimated for each confidence level.

Firstly, we focus on the relationship between confidence level and overall ACP, shown in Figure 5.3. As expected, overall ACP decreases as the confidence level increases. This happens because higher confidence levels result in higher q_j values, which according to the method's decision rule will inevitably lead to less accepted parcel classification decisions. With the low confidence level of 50%, almost all parcels can be classified automatically (97.6%). Using such a low λ is not recommendable since the goal of the method is to eliminate parcels that are responsible for bad classification performances. If a 100% accurate classification is demanded, only 5.5% of all parcels can be classified in an automatic way. We do not possess information about the accuracy of the traditional CAPI process, but we can state that it is most likely

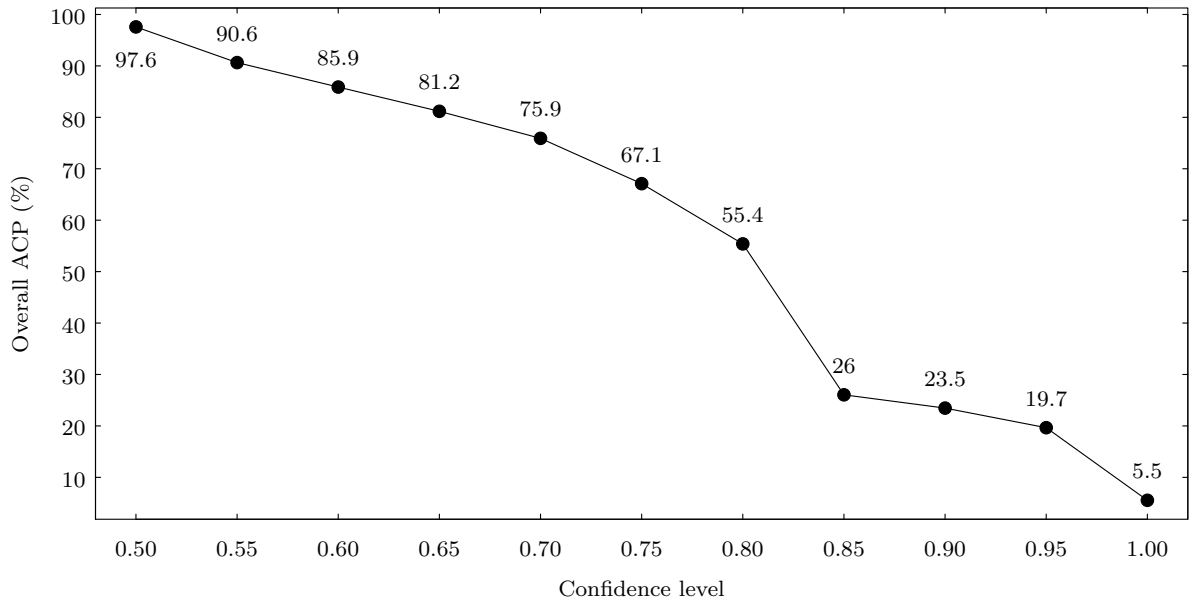


Figure 5.3: Overall proportion of agriculture parcels that can be classified automatically with different classification confidence level values. At a low value of $\lambda = 50\%$, almost all parcels are classified in an automatic way. On the other hand, a completely accurate classification only applies to 5.5% of classified parcels. A trade-off between reliability and number of automatic classifications with respective time and money savings has to be undertaken.

not completely accurate (Congalton, 1991). It should therefore not be expected from a remote sensing application that it correctly classifies all given objects. An interesting confidence level is 80%, which leads to classification errors below 20% and classifies more than 50% of all parcels in an automatic way. Note the large drop in overall ACP that takes place between 80% and 85% of confidence, falling from over 55% to only 26%, respectively.

Secondly, the relationship between confidence level and ACP for each individual land cover class was analyzed in Figure 5.4. The general trend visible in Figure 5.3 is obviously also present here: higher λ values lead to lower ACP values. The only exceptions are Rice and Poor grassland, which remain constant at 100% and 2.3% of classified parcels irrespective of λ. Over 99% of parcels classified as Maize are accepted with 60%, 70%, 80% and 90% of confidence, with a small drop for $\lambda = 95\%$. Focus should be put on Permanent grassland, since this class accounts for more than half of the total studied area. ACP is high and moderate for $\lambda \leq 80\%$ but falls to near-zero 0.3% of classified parcels for $\lambda = 90\%$. This explains the large drop of the overall ACP from 55.4% to 23.5%, for confidence of 80% and 90%, respectively. It is noteworthy that many classes show a near-100% ACP for $\lambda = 60\%$ and even 70%, explaining the relatively high overall ACP for these confidence levels (85.9% and 75.9%, respectively). On the contrary, some classes reveal low classification proportions even at low confidence levels, such as Forage crops, Fallow, Poor grassland and Non used area. In these classes, ACP drops to near-zero values for

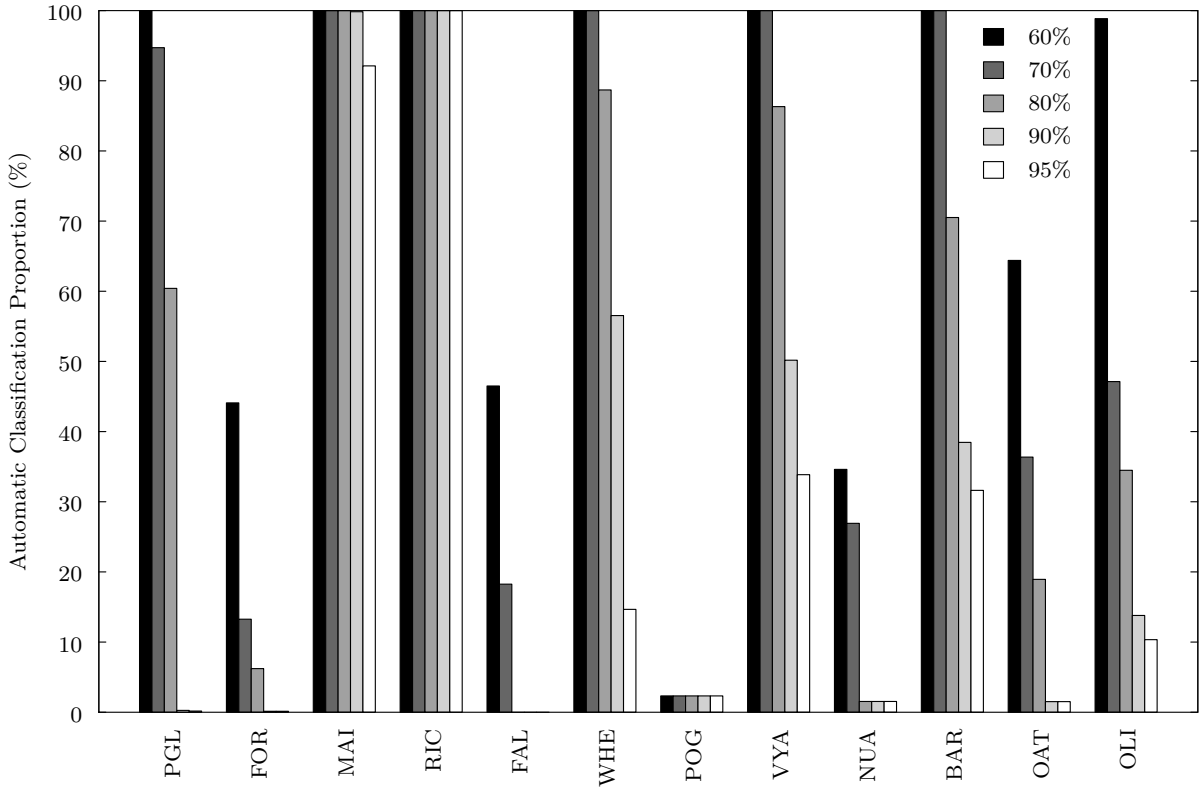


Figure 5.4: Proportion of agriculture parcels that can be classified automatically with different classification confidence level values in each land cover class. Note the very distinct behavior of the different classes. It is notable that many classes show a near-100% ACP for $\lambda = 60\%$ and 70% while other classes show low classification proportions even at low confidence levels.

higher confidences. Automatic classification is not able to handle these classes in a satisfactory way. As discussed before, the reason behind this behavior is that these classes are difficult to discriminate from one another based on their multitemporal spectral signature.

Large amounts of time and money are spent to perform the required checks in each control campaign. Therefore, the overall goal of the method is to reduce OTSC costs and completion time. Both can be reduced as the number of automatically classified parcels that do not require the intervention of a photo interpreter increases. According to the presented results, low confidence levels can save more time and money than high confidence levels. On the other hand, low confidence levels increase the number of undesirable misclassifications. A careful trade-off between time and money savings and the reliability of the control system must therefore be carried out by the decision-maker before using this method in an operational context. In conjunction with the portuguese decision-maker IFAP, we chose a confidence level of 80% to carry out accuracy assessment of the automatization method. An overall UA of 80% ensures a reasonable system reliability and enables the automatic control of more than half (55.4%) of the required parcels, also cutting by more than half time and money expenditures. As for Maize, Rice, Wheat and Vineyard, even more than 85% of all parcels can be controlled by this method using the selected λ. Moreover, choosing a confidence level higher than 80% would almost entirely prevent classification of the most populated class Permanent grassland, which is a very important class due to its frequency on the ground.

5.3.2 Automatic classification assessment

The method was calibrated and assessed with the selected confidence level of 80%. A corresponding error matrix was generated and the usual accuracy statistics were estimated, shown in Table 5.3. A summary of q_j and ACP values for that confidence level are also provided in Table 5.4. The OA of this classification was 84.1% which we consider as a good performance. This result constitutes another advantage of using $\lambda = 80\%$ for automatic classification. All classes reveal UA $\geq \lambda$, proving that the method was calibrated successfully.

The crop Rice yielded the best results with both user and producer accuracies of approximately 98%. Other authors have also found this crop to be the best performing, attributing this efficiency to the fact that this crop grows in flooded fields, which are very distinguishable due to the effect of water in the NIR and Short Wave Infrared (SWIR) spectral regions (Peña Barragán et al., 2011). Maize also shows excellent results, with a UA = 89.9% substantially higher than the established minimum of 80%. Maize and Rice represent the best performing classes, given its very satisfactory accuracies along with 100% of automatically classified parcels in both classes (Table 5.4). Permanent grassland shows a very high PA of 98.5% but a UA equaling the confidence level 80%, which is because this crop is highly confused with many other classes as revealed by the error matrix. The same situation happens with Vineyard, which also shows a very high PA and confusion with other classes. Our best performing crops Maize, Rice and Vineyard are among the most accurate classes reported by other studies such as the work

Table 5.3: Error matrix for automatic land cover classifications with confidence level $\lambda = 80\%$ and accuracy statistics UA and PA. OA for this classification was 84.1%. Note that the chosen confidence level is reflected in the UA $\geq 80\%$ yielded in each class.

	Reference land cover class												**UA (%)**
	PGL	FOR	MAI	RIC	FAL	WHE	POG	VYA	NUA	BAR	OAT	OLI	
PGL	2563	264	0	0	138	0	129	8	77	0	9	15	80
FOR	7	73	0	0	5	2	0	0	0	1	3	0	80.2
MAI	13	32	1141	16	42	9	0	4	5	2	5	0	89.9
RIC	1	0	7	772	1	0	0	0	4	0	0	0	98.3
FAL	0	0	0	0	0	0	0	0	0	0	0	0	n.d.
WHE	2	19	2	0	12	402	0	0	2	35	27	1	80.1
POG	0	0	0	0	0	0	1	0	0	0	0	0	100
VYA	15	22	4	1	41	1	0	394	2	0	0	12	80.1
NUA	0	0	0	0	0	0	0	0	2	0	0	0	100
BAR	0	6	0	0	1	20	0	0	0	132	6	0	80
OAT	0	4	0	0	1	0	0	0	0	0	20	0	80
OLI	2	1	0	0	3	0	0	0	0	0	0	24	80
PA (%)	98.5	17.3	98.9	97.8	0	92.6	0.8	97	2.2	77.6	28.6	46.2	

of Peña Barragán et al. (2011), showing that our results are consistent with the literature. Poor grassland and Non used area have completely accurate classifications in terms of UA, but a more careful analysis of these results shows that only 1 and 2 parcels were classified as Poor grassland and Non used area, respectively. Also, PA values of those classes are near-zero, 0.8% and 2.2% respectively, revealing that the classifier is not able to deal with those classes in a satisfactory manner. Results from Table 5.4 further sustain the conclusion that Poor grassland and Non used area are rather insignificant in this classification, as suggested by the small proportion of automatically classified parcels (ACP values of respectively 2.3% and 1.5%). The situation in Fallow is even worse, with no parcel being automatically assigned to this class. This led to PA and ACP of 0 and an undefined q_j. An inspection of the error matrix shows that those classes are mainly confused with Permanent grassland by the classifier. In fact, it was previously stated that those 4 classes show similar average spectral signatures which lead to the inevitable confusion between classes. Based on this discussion, we recommend that no automatic decisions should be done in classes Fallow, Poor grassland and Non used area. They actually represent very similar land cover types on the ground, constituted by a ill-defined mixture of different vegetation types or even bare soil. As for the winter cereal crops, namely Wheat, Barley and Oat, as expected some confusion happens between the respective classes. UA rounds 80% for all 3 classes. Wheat and Barley show high PA, indicating that the classifier performs well in separating those classes, while it is relatively low for Oat.

The following example shows the reason for the focus on UA rather than on PA. PA and UA for Permanent grassland were respectively 98.5% and 80%. These values indicate that

Table 5.4: Estimates of q_j and ACP, both with a confidence level $\lambda = 80\%$.

Class	PGL	FOR	MAI	RIC	FAL	WHE	POG	VYA	NUA	BAR	OAT	OLI
$\mathbf{q_j}$	0.722	0.686	0.239	0.257	n.d.	0.407	0.636	0.426	0.733	0.487	0.683	0.492
ACP (%)	60.4	6.2	100	100	0	88.7	2.3	86.3	1.5	70.5	18.9	34.5

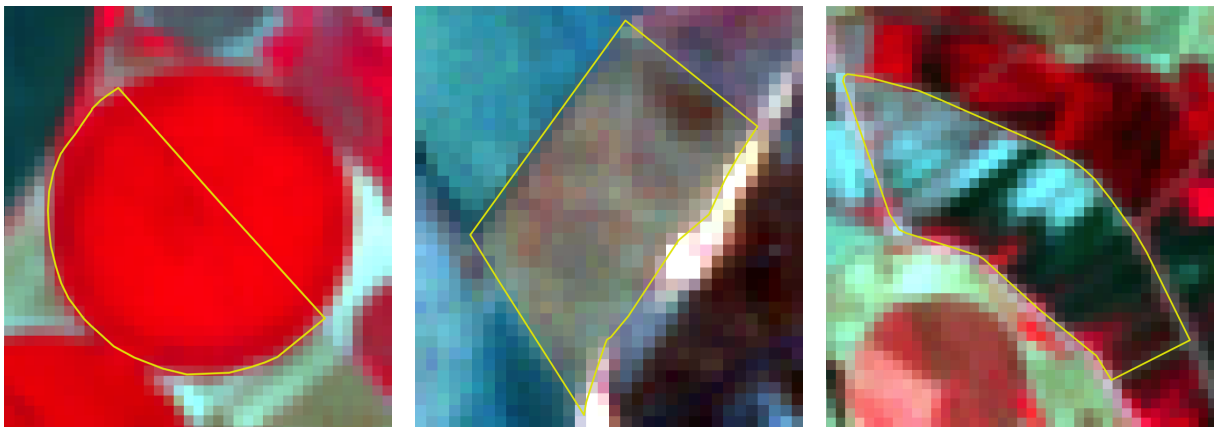
98.5% of the Permanent grassland parcels on the ground were correctly identified as Permanent grassland, but only 80% of the parcels that were assigned Permanent grassland by the classifier were actually Permanent grassland. Considering the objectives of CAPI and subsidy control in the context of the CAP, it becomes obvious that the control system needs to maximize UA and not PA. Therefore, there is no problem in the low PA values that some classes exhibit, as long as a high UA is guaranteed, which is taken care of by the method's confidence level. Also regarding PA, one thing that can be noticed from simultaneous analysis of Table 5.3 and 5.4 is that PA is correlated with ACP. High PA values tend to correspond to high ACP values. This behavior makes sense, since in some way both can be seen as a measure of how well the classifier can handle classification of a given land cover class.

Generally speaking, it was expected from average spectral signatures analysis that the classes would exhibit very divergent classification performances. Considering the discussed results for the confidence level of 80%, we recommend that all classification decisions in classes Maize and Rice are accepted by the decision-maker, while rejecting all decisions in classes Fallow, Poor grassland and Non used area. The remaining classes should be handled according to the method's decision rule, using estimated q_j values.

5.3.3 An application example

A practical method application example is provided for the purpose of giving the reader a better understanding of how the method works in the operational context. Classification results with two different confidence levels are discussed in this example: 80% and 95%. Figure 5.5 shows false color composites of three selected parcels with known land cover. The corresponding classification results can be found in Table 5.5.

Parcel (a) shows a homogeneous maize cropping with very high spectral response, revealed by the bright red color in the false color composite. This led to a clear classification of the parcel as Maize with a high posterior probability $P(\omega_j|\mathbf{x})$ of 0.95. This result was accepted



(a) Maize (41.1 ha) (b) Permanent grassland (35.9 ha) (c) Rice (44.8 ha)

Figure 5.5: False color composite (R = NIR, G = Red, B = Green) at July 8, 2005 of three selected parcels with different known land cover classes for exemplification of the method's application phase (see Table 5.5).

Table 5.5: Automatic classification results for parcels represented in Figure 5.5. q_j values and corresponding results are with $\lambda = 80\%$ and $\lambda = 95\%$.

Parcel	**True class**	**Classification output**		$\mathbf{q_{j,80}}$	$\mathbf{q_{j,95}}$	$\mathbf{Result_{80}}$	$\mathbf{Result_{95}}$
		Decision ω_j	$P(\omega_j\|\mathbf{x})$				
(a)	MAI	MAI	0.950	0.239	0.439	Accepted	Accepted
(b)	PGL	FOR	0.647	0.686	0.831	Rejected	Rejected
(c)	RIC	MAI	0.389	0.239	0.439	Accepted	Rejected

for both confidence levels and the parcel was classified automatically, since $P(\omega_j|\mathbf{x})$ was higher than both $q_{j,80}$ and $q_{j,95}$. Parcel (b) is covered by Permanent grassland and shows a rather heterogenous crop pattern. In the classification process, it was assigned to the incorrect class Forage crops with posterior probability 0.647. However, the method correctly rejected that classification and decided not to classify the parcel in an automatic way, for both confidence levels. For 80% confidence, it was a close rejection, given that $P(\omega_j|\mathbf{x})$ almost matched the minimum of 0.686. As for 95%, the classification is clearly rejected, which makes sense, since demanding a higher confidence in the classification decisions increases the threshold to accept classifications, ultimately resulting in more rejected parcels. This parcel is a perfect example to demonstrate the usefulness of the proposed method, rejecting parcels that can not be classified with the required certainty. The third example is parcel (c) which corresponds to Rice, revealing a very heterogenous behavior. This parcel was also assigned to the wrong class by the classifier, in this case to Maize with an associated probability of 0.389. The low q_j for Maize allows for misclassifications when parcels are classified as Maize even with relatively low posterior probabilities, which happened in this case. The classification was wrongly accepted by the method with confidence level 80% and correctly rejected with confidence 95%. This example clearly shows the positive effect a higher confidence level has on the accuracy of an automatic classification decision, with the drawback of automatically classifying a smaller proportion of the total number of parcels.

6 Conclusions and final remarks

In this study, we proposed a method for automating the CAPI task in the context of CAP-based subsidy control. Parcel-based land cover classification in the Ribatejo region was carried out using SVM and KNN classifiers, validated control data from the portuguese decision-maker IFAP and an atmospherically corrected Landsat 7 ETM+ time series. Performance of remotely sensed data dimensionality reduction was assessed in order to remove possible redundancy in the data. The method was build upon the assumption that it is not viable to classify all agricultural parcels which require OTSC. Significant classification errors would be committed in many land cover classes, compromising the control process. We introduced a confidence level in order to enable adjustment of the control system's reliability by the decision-maker. In simple terms, this confidence level selects a subset of parcels requiring OTSC which are classified automatically with the desired overall reliability. The following conclusions can be drawn from the results of this work:

1. Dimensionality of the original multispectral and multitemporal feature space was successfully reduced from 36 to 12 without compromising classification accuracy of the SVM classifier. The variable selection indicates that Red and NIR bands are the most relevant.

2. SVM classifier performed better than KNN using both the original and the dimensionality reduced datasets. However, OA of all combinations of classifiers and datasets was substantially below the recommended 85%, justifying the need for the proposed CAPI automatization method.

3. Low confidence levels of the method lead to high proportions of automatically classified parcels, therefore reducing CAPI costs and completion time. However, a low confidence level increases the risk of misclassifications, compromising the reliability of the control system. A confidence level of 80% was found to be the right balance between these two contrasting criteria.

4. With confidence level 80%, over 55% of all parcels can be controlled automatically, replacing the traditional CAPI process. Of course, the remaining parcels must be subjected to CAPI. Moreover, 85% of all parcels classified as Maize, Rice, Wheat or Vineyard can be classified automatically. The OA of classification with 80% confidence was 84% over all 12 classes.

5. Again, for an 80% confidence level, all classification decisions in classes Maize and Rice can be accepted by the decision-maker. On the other hand, all decisions in classes Fallow, Poor grassland and Non used area should be rejected due to extreme confusion between these classes. The remaining classes should be handled according to the method's decision rule.

6. The goal of developing a simple, cost-effective and reproducible method to automatize the CAPI process in an operational context was successfully achieved. CAPI costs and completion time will be significantly reduced once the method is adopted by the decision-maker. Furthermore, we hope our method is seen as an incentive to substantially raise the low proportion of parcels that is annually controlled via remote sensing in the portuguese case, slightly over 25% in the last years.

7. With this study we demonstrated the potential and the feasibility of automatic land cover control via remote sensing over the large Mediterranean region Ribatejo.

Despite the satisfactory outcomes of this study, there were also some constraints which may have affected classification performance. We believe that the achieved results, namely ACP and OA, could be further improved. These constraints included:

1. A rather large number of agricultural parcels was affected by the SLC-off issue, reducing the amount of available spectral information lying within these parcels.

2. The relatively low spatial resolution of ETM+ (multispectral 30 m) when compared to other HR sensors limited the analysis of smaller agricultural parcels.

3. Not all available multispectral images matched the established optimal image acquisition periods used by IFAP.

4. The number of elements in each land cover class is relatively unbalanced, with the most populated class holding 16 times more parcels than the less represented class. In terms of area, this difference is even greater, with Permanent grassland holding 56% of the entire study area while Olive grove represents only 0.7%.

Some of these constraints lead us to suggest directions for further research, with the goal of improving the proposed method in terms of performance and usability:

1. The portuguese agricultural landscape reveals very distinct patterns across the country. For instance, in regions such as Northern Portugal where intercropping prevails and extremely small agricultural parcels are commonly found, the application of the method with the used spatial resolution is nearly impossible. The usage of an alternative imagery source with higher spatial resolution for calibration and application of the method could be explored to partially overcome this problem, enabling the automatic control of smaller parcels. SPOT-5 (multispectral 10 m), SPOT-6 and SPOT-7 (both multispectral 6 m), as well as the upcoming Sentinel-2 mission (multispectral 10 m, 20 m and 60 m) are best suited for providing alternative HR imagery.

2. Our approach is conducive to a muti-sensor analysis, since each parcel is described by its average reflectance. Landsat 7 data can therefore be integrated with data obtained from other satellites. A multi-sensor approach has two advantages: firstly, having multiple sensors is equivalent to having a more populated range of image acquisition dates to choose

from. This means that images can be chosen to optimally follow the crops' development cycle in both the method's calibration and application phase. Secondly, this approach facilitates coping with missing data due to clouds or other constraints. Sentinel-2 could play an important role in this process due to its high revisit rate below 5 days.

3. We assume that the obtained results are valid exclusively for the selected study area in Ribatejo. Therefore, the ability to generalize the calibrated method and corresponding classification results to other Mediterranean regions must be studied in order to extend it to other areas with similar climatic conditions.

4. The selection of image acquisition dates is done so to optimally follow the typical crop development cycle throughout the year, with application phase images supposed to match calibration phase images as well as possible. The reason behind this is the assumption that crops' multitemporal spectral signatures show near-constant behavior from year to year. However, this assumption can be violated by atypical crop development due to unusual climatic conditions in a given year. MARS releases monthly crop monitoring bulletins that offer analyses and information on important crop types growth conditions and yield forecast at EU28 level (MARS, 2014). The information provided in these reports could be used in further research to adjust the image acquisition dates in a given year, in order to follow the actual crops' development cycle in that year.

We finish this study with practical recommendations on how to apply the method developed and presented in this work in the near future. Landsat 8 imagery is already available, offering significant improvements over previous Landsat missions in both data quality and spectral coverage (Irons et al., 2012). In this work, we used data from the same year 2005 to both calibrate and apply automatic classification, which is of course not the objective in the operational context of the portuguese decision-maker IFAP. For the purpose of calibrating automatic control for the next control period in 2015, we recommend the simultaneous usage of Landsat 8 and control 2014 data in a selected control area. In the application phase taking place in 2015, Landsat 8 imagery can then be used to automatize the CAPI process in accordance with this study. The present recommendation can obviously be changed according to results of further research. For instance, the usage of other higher resolution sensors in the future, providing the advantages described earlier, should be taken into consideration. The following assumptions are made in the current version of the method: 1) acquisition dates of calibration and application image time series should be as identical as possible, 2) the calibration and application image time series should use the same method for atmospheric effect correction, which is mandatory, and 3) the application phase should take place in the area where the method was calibrated.

Bibliography

APA. Atlas do Ambiente. URL http://sniamb.apambiente.pt/atlas/. [Accessed 07/05/14].

A. Belousov, S. Verzakov, and J. von Frese. A flexible classification approach with optimal generalisation performance: support vector machines. *Chemometrics and Intelligent Laboratory Systems*, 64(1):15–25, 2002.

R. Bivand. maptools package for R, version 0.8–29, 2014.

R. Bivand, T. Keitt, B. Rowlingson, E. Pebesma, M. Sumner, R. Hijmans, and E. Rouault. rgdal package for R, version 0.8–16, 2014a.

R. Bivand, C. Rundel, E. Pebesma, and K. O. Hufthammer. rgeos package for R, version 0.3–4, 2014b.

X. Blaes, L. Vanhalle, and P. Defourny. Efficiency of crop identification based on optical and SAR image time series. *Remote Sensing of Environment*, 96(3-4):352–365, 2005.

A. Brenning. Benchmarking classifiers to optimally integrate terrain analysis and multispectral remote sensing in automatic rock glacier detection. *Remote Sensing of Environment*, 113(1):239–247, 2009.

C. J. C. Burges. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2:121–167, 1998.

J. Cadima, J. O. Cerdeira, P. D. Silva, and M. Minhoto. subselect package for R, version 0.12–3, 2013.

Câmara Municipal de Santarém. Caracterização Ambiental. URL http://www.cm-santarem.pt/concelho/caracterizacaodoconcelho/Paginas/ambiental.aspx. [Accessed 07/05/14].

V. M. P. Carmona. *Relatório integrador da actividade profissional. Os sistemas de informação geográfica (SIG) no controlo de superfícies agrícolas em Portugal Continental.* Master's thesis, Instituto Superior de Agronomia – Universidade Técnica de Lisboa, 2012.

V. M. P. Carmona. Personal comunication, 2014.

J. a. M. Carreiras, J. M. Pereira, M. L. Campagnolo, and Y. E. Shimabukuro. Assessing the extent of agriculture/pasture and secondary succession forest in the Brazilian Legal Amazon using SPOT VEGETATION data. *Remote Sensing of Environment*, 101(3):283–298, 2006.

L. Carvalho. Conceito de Terroir na Região Tejo: Castas e Solos, 2010. URL http://www.cvrtejo.pt/dlds/Apresentacao_Lus_Carvalho-O_Conceito_de_Terroir_na_Regio_Tejo-Castas_e_Solos.pdf.

I. L. Castillejo-González, F. López-Granados, A. García-Ferrer, J. M. Peña Barragán, M. Jurado-Expósito, M. S. de la Orden, and M. González-Audicana. Object- and pixel-based analysis for mapping crops and their agro-environmental associated measures using QuickBird imagery. *Computers and Electronics in Agriculture*, 68(2): 207–215, 2009.

G. Chander, B. L. Markham, and D. L. Helder. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment*, 113(5):893–903, 2009.

C.-C. Chang and C.-J. Lin. LIBSVM: A Library for Support Vector Machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3):1–27, 2011.

R. G. Congalton. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37(1):35–46, 1991.

C. Conrad, S. Fritsch, J. Zeidler, G. Rücker, and S. Dech. Per-Field Irrigated Crop Classification in Arid Central Asia Using SPOT and ASTER Data. *Remote Sensing*, 2(4):1035–1056, 2010.

B. B. Damodaran and R. R. Nidamanuri. Assessment of the impact of dimensionality reduction methods on information classes and classifiers for hyperspectral image classification by multiple classifier system. *Advances in Space Research*, 53(12):1720–1734, 2014.

R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification (2nd edition)*. John Wiley, New York, 2000.

EC. Commission Regulation No 1307/2013. 2008(1307):608–670, 2013a.

EC. The common agricultural policy (CAP) and agriculture in Europe – frequently asked questions, 2013b. URL http://ec.europa.eu/agriculture/faq/index_en.htm#2. [Accessed 07/28/14].

M. El Hajj, A. Bégué, S. Guillaume, and J.-F. Martiné. Integrating SPOT-5 time series, crop growth modeling and expert knowledge for monitoring agricultural practices – The case of sugarcane harvest on Reunion Island. *Remote Sensing of Environment*, 113(10):2052–2061, 2009.

FÖMI. CwRS – Control with Remote Sensing of Area-based Subsidies, 2013. URL http://www.fomi.hu/portal_en/index.php/products-and-services/cwrs-control-with-remote-sensing-of-area-based-subsidies. [Accessed 07/24/14].

G. M. Foody. Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80(1): 185–201, 2002.

R. J. Hijmans. raster package for R, version 2.2–31, 2014.

C. Huang, L. S. Davis, and J. R. G. Townshend. An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, 23(4):725–749, 2002.

C. Huang, K. Song, S. Kim, J. R. Townshend, P. Davis, J. G. Masek, and S. N. Goward. Use of a dark object concept and support vector machines to automate forest cover change analysis. *Remote Sensing of Environment*, 112(3):970–985, 2008.

IES. Earth Observation data supporting the CAP implementation, 2013. URL http://ies.jrc.ec.europa.eu/our-activities/support-for-eu-policies/control-with-remote-sensing.html. [Accessed 07/21/14].

S. B. Imandoust and M. Bolandraftar. Application of K-Nearest Neighbor (KNN) Approach for Predicting Economic Events: Theoretical Background. *Int. Journal of Engineering Research and Applications*, 3(5):605–610, 2013.

Infovini. Regiões Vitivinícolas, 2014. URL http://www.infovini.com/classic/pagina.php?codPagina=10&regiao=8&flash=1. [Accessed 07/05/14].

IPMA. Clima de Portugal Continental, 2014. URL http://www.ipma.pt/pt/educativa/tempo.clima/index.jsp?page=clima.pt.xml. [Accessed 07/05/14].

J. R. Irons, J. L. Dwyer, and J. A. Barsi. The next Landsat satellite: The Landsat Data Continuity Mission. *Remote Sensing of Environment*, 122:11–21, 2012.

I. T. Jolliffe. *Principal component analysis (2nd edition)*. Springer, New York, 2002.

JRC. Summary Report of 2012 CwRS Image Acquisition Campaign. Technical report, Joint Research Centre, Luxembourg: Publications Office of the European Union, 2012.

T. Kavzoglu and I. Colkesen. A kernel functions analysis for support vector machines for land cover classification. *International Journal of Applied Earth Observation and Geoinformation*, 11(5):352–359, 2009.

P. B. C. Leite, R. Q. Feitosa, A. R. Formaggio, G. A. O. P. da Costa, K. Pakzad, and I. D. Sanches. Hidden Markov Models for crop recognition in remote sensing image sequences. *Pattern Recognition Letters*, 32(1): 19–26, 2011.

H.-T. Lin, C.-J. Lin, and R. C. Weng. A note on Platt's probabilistic outputs for support vector machines. *Machine Learning*, 68(3):267–276, 2007.

P. Loudjani. G-tech supports a common agriculture policy in Europe. *Geospatial World*, 2013. URL http://geospatialworld.net/Magazine/MArticleView.aspx?aid=30677.

A. R. S. Marçal, J. S. Borges, J. A. Gomes, and J. F. Pinto Da Costa. Land cover update by supervised classification of segmented ASTER images. *International Journal of Remote Sensing*, 26(7):1347–1362, 2005.

MARS. Geomatics in support of the Common Agricultural Policy – Proceedings of the 14th MARS PAC Annual Conference, 2008.

MARS. MARS Bulletins for Europe, 2014. URL http://mars.jrc.ec.europa.eu/mars/About-us/AGRI4CAST/MARS-Bulletins-for-Europe. [Accessed 10/03/14].

D. Meyer, E. Dimitriadou, K. Hornik, A. Weingessel, F. Leisch, C.-C. Chang, and Chih-Chen Lin. e1071 package for R, version 1.6–4, 2014.

G. Mountrakis, J. Im, and C. Ogole. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3):247–259, 2011.

J. M. Peña Barragán, M. K. Ngugi, R. E. Plant, and J. Six. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sensing of Environment*, 115(6):1301–1316, 2011.

E. Pebesma, R. Bivand, B. Rowlingson, and V. Gomez-Rubio. sp package for R, version 1.0–14, 2013.

R Development Core Team. The Comprehensive R Archive Network, 2014. URL http://cran.r-project.org/.

RStudio. RStudio – Open source and enterprise-ready professional software for R, 2014. URL http://www.rstudio.com/.

V. Sagris, P. Wojda, P. Milenov, and W. Devos. The harmonised data model for assessing Land Parcel Identification Systems compliance with requirements of direct aid and agri-environmental schemes of the CAP. *Journal of environmental management*, 118:40–8, 2013.

K. Schliep and K. Hechenbichler. kknn package for R, version 1.2–5, 2014.

M. Singh, A. Marchis, and E. Capri. Greening, new frontiers for research and employment in the agro-food sector. *The Science of the total environment*, 472:437–43, 2014.

C. Song, C. E. Woodcock, K. C. Seto, M. P. Lenney, and S. A. Macomber. Classification and Change Detection Using Landsat TM Data. *Remote Sensing of Environment*, 75(2):230–244, 2001.

B. M. Steele. Maximum posterior probability estimators of map accuracy. *Remote Sensing of Environment*, 99 (3):254–270, 2005.

S. Urbanek. proj4 package for R, version 1.0–8, 2014.

USGS. Landsat Surface Reflectance Climate Data Records, 2014a. URL http://pubs.usgs.gov/fs/2013/3117/pdf/fs2013-3117.pdf.

USGS. Product Guide – Landsat Climate Data Record (CDR) Surface Reflectance, 2014b. URL http://landsat.usgs.gov/documents/cdr_sr_product_guide.pdf.

K. Usha and B. Singh. Potential applications of remote sensing in horticulture – A review. *Scientia Horticulturae*, 153:71–83, 2013.

S. M. Vicente-Serrano, F. Pérez-Cabello, and T. Lasanta. Assessment of radiometric correction techniques in analyzing vegetation variability and change using time series of Landsat images. *Remote Sensing of Environment*, 112(10):3916–3934, 2008.

B. D. Wardlow, S. L. Egbert, and J. H. Kastens. Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains. *Remote Sensing of Environment*, 108(3):290–310, 2007.

Wikicap. Computer Assisted Photo-interpretation, 2011a. URL http://marswiki.jrc.ec.europa.eu/wikicap/index.php/Computer_Assisted_Photo-interpretation_(CAPI). [Accessed 07/17/14].

Wikicap. Image Processing, 2011b. URL http://marswiki.jrc.ec.europa.eu/wikicap/index.php/Image_Processing. [Accessed 07/17/14].

Wikicap. CwRS, 2012. URL http://marswiki.jrc.ec.europa.eu/wikicap/index.php/CwRS. [Accessed 07/03/14].

Wikicap. Principles of Control with Remote Sensing and possible strategies, 2013. URL http://marswiki.jrc.ec.europa.eu/wikicap/index.php/Principles_of_Control_with_Remote_Sensing_and_possible_strategies. [Accessed 07/17/14].

C. Yang, J. H. Everitt, and D. Murden. Evaluating high resolution SPOT 5 satellite imagery for crop identification. *Computers and Electronics in Agriculture*, 75(2):347–354, 2011.