THE CONTRIBUTION OF MULTITEMPORAL INFORMATION FROM MULTISPECTRAL SATELLITE IMAGES FOR AUTOMATIC LAND COVER CLASSIFICATION AT THE NATIONAL SCALE

by

Hugo Miguel Saiote Carrão

Dissertation submitted to the Instituto Superior de Estatística e Gestão de Informação da Universidade Nova de Lisboa in partial fulfillment of the requirements for the Degree of

Doutor em Gestão de Informação – Sistemas de Informação Geográfica

(Doctor of Philosophy in Information Management – Geographic Information Systems)

Instituto Superior de Estatística e Gestão de Informação

da

Universidade Nova de Lisboa

2009

THE CONTRIBUTION OF MULTITEMPORAL INFORMATION FROM MULTISPECTRAL SATELLITE IMAGES FOR AUTOMATIC LAND COVER CLASSIFICATION AT THE NATIONAL SCALE

by

Hugo Miguel Saiote Carrão

Dissertation supervised by

Professor Mário Caetano

Professor Paulo Gonçalves

Instituto Superior de Estatística e Gestão de Informação

da

Universidade Nova de Lisboa

2009

DEDICATION

To my parents and brother.

ACKNOWLEDGMENTS

A few lines are too short to make a complete account of my deep appreciation for my supervisors. As most of the present PhD thesis was completed in the Portuguese Geographic Institute (IGP), under the supervision of Professor Mário Caetano, I wish to thank him for his unflattering trust and constant encouragements, which have been essential to our success through the last four years. Thank you for being an enthusiastic and sagacious individual who provided a steady support in the most difficult times. My apprentice pursuit into the world of science would not have been the same without Professor Paulo Gonçalves's approach to research and science. I want to extend my appreciation for his patience, for his full devotion, for his great understanding, and for the lessons he gave me on the importance of details for the success of engineering and scientific projects.

I am indebted to my friends and colleagues from IGP for providing me a good environment in which to work. I am grateful to António Araújo, António Nunes, Conceição Pereira and Vasco Nunes for bearing my recurrent bad temper and irascibility trough those last years.

Sincere thanks are extended to Professors Fernando Bação and Pedro Coelho, on behalf of making many valuable suggestions which helped improve this thesis.

I wish to express my cordial appreciation to Professors Florence Forbes and Pascale Primet for accepting me as a member of their research teams during my PhD work. Sincere thanks are extended to teams' members and special friends from Lyon, in particular Dinil Divakaran, Suleyman Baykut and Marianne Schaedel.

I wish to thank to my friends Sérgio Freire, Marco Freire, André Oliveira, Ricardo Armas, André Pinheiro, Susana Dias, Helena Barahona, Inês Pais and Margarida Ferreira for all the emotional support, camaraderie, entertainment, and caring they provided.

I would like to express my heartiest thanks to Marta Jaco and Tatiana Plantier for being with me during the happiest and saddest times, and for their love, tenderness and trust.

I wish to thank to my brother Ricardo, my grandfather José Saiote, and my uncles Fernando Saiote and António Saiote for their encouragements throughout my graduate work.

I owe my parents, Mariana Saiote and Armando Carrão, much of what I have become. I thank them for their love, their support, and their confidence throughout the past thirty-one years. They taught me to value honesty and humility above all other virtues. Thank you for your unwavering support.

A big thank you to you all!

The research presented in this dissertation was partially funded and/or supported by the following institutions:

- *Fundação para a Ciência e Tecnologia* (FCT) – SFRH/BD/18447/2004;
- *Programa Operacional Ciência e Inovação* 2010 (POCI 2010);
- European Union (EU);
- European Space Agency (ESA);
- *Instituto Geográfico Português* (IGP);
- *Institut National de Recherche en Informatique et Automatique* (INRIA) *Rhône-Alpes*;
- *École Normale Supérieure de Lyon* (ENS) - *Laboratoire de l'Informatique du Parallélisme* (RESO – INRIA);
- Luso-French Integrated University Actions – 2007-2008;
- Research Centre for Statistics and Information Management (CEGI).
- *Fundação Luso-Americana* (FLAD);
- *Fundação Calouste Gulbenkian.*

RESUMO

As tecnologias de informação associadas à detecção remota estão em constante evolução, tal que, actualmente, os novos satélites permitem a recolha diária de imagens da superfície terrestre. É inquestionável, que a tendência para uma observação temporal contínua da superfície terrestre irá alargar o âmbito de actuação das actividades relacionadas com a detecção remota. Torna-se inevitável, com o crescimento abrupto da informação temporal disponível, o desenvolvimento de abordagens que combinem o processamento digital de imagens de satélite multi-espectrais com a análise tradicional de séries temporais, para a caracterização da distribuição da ocupação do solo e da sua monitorização. No entanto, métodos quantitativos que demonstrem a proficiência de séries tridimensionais de imagens de satélite (i.e. especial, espectral e temporal), para a classificação automática da ocupação do solo e análise da dinâmica da ocupação do solo, não foram todavia explorados. Nesta dissertação investiga-se a utilidade de séries temporais de imagens de satélite multi-espectrais para a caracterização regular da ocupação do solo a uma escala nacional. Este estudo é realizado no território de Portugal Continental e explora imagens de satélite adquiridas pelo Moderate Resolution Imaging Spectroradiometer (MODIS) e MEdium Resolution Imaging Spectrometer (MERIS). Especificamente, focamo-nos na análise da contribuição da informação temporal de imagens de satélite multi-espectrais para a discriminação automática de classes de ocupação do solo. Os resultados mostram que a informação multi-espectral contribui de forma mais significativa do que a informação multitemporal para a classificação automática de tipos de ocupação do solo. Na sequência, revemos algumas dos pontos mais importantes de qualquer protocolo de produção automática de cartografia de ocupação do solo a partir de imagens de satélite. Propomos uma abordagem metodológica para a produção automática de mapas de ocupação do solo a partir de séries temporais de imagens de satélite e sua validação. Finalmente, desenvolvemos um modelo harmónico não linear para o ajuste temporal de reflectâncias e índices de vegetação derivados de imagens de satélite, para diversas classes de ocupação do solo. A informação multitemporal sintetizada pelo modelo é suficientemente abrangente para descrever as principais características da ocupação do solo e prever a evolução temporal dos seus indivíduos.

Palavras-Chave: Detecção remota, séries temporais, imagens multi-espectrais, ocupação do solo, análise de separabilidade, classificação, predição, dados de referência, avaliação da exactidão.

ABSTRACT

Imaging and sensing technologies are constantly evolving so that, now, the latest generations of satellites commonly provide with Earth's surface snapshots at very short sampling periods (i.e. daily images). It is unquestionable that this tendency towards continuous time observation will broaden up the scope of remotely sensed activities. Inevitable also, such increasing amount of information will prompt methodological approaches that combine digital image processing techniques with time series analysis for the characterization of land cover distribution and monitoring of its dynamics on a frequent basis. Nonetheless, quantitative analyses that convey the proficiency of three-dimensional satellite images data sets (i.e. spatial, spectral and temporal) for the automatic mapping of land cover and land cover time evolution have not been thoroughly explored. In this dissertation, we investigate the usefulness of multispectral time series sets of medium spatial resolution satellite images for the regular land cover characterization at the national scale. This study is carried out on the territory of Continental Portugal and exploits satellite images acquired by the Moderate Resolution Imaging Spectroradiometer (MODIS) and MEdium Resolution Imaging Spectrometer (MERIS). In detail, we first focus on the analysis of the contribution of multitemporal information from multispectral satellite images for the automatic land cover classes' discrimination. The outcomes show that multispectral information contributes more significantly than multitemporal information for the automatic classification of land cover types. In the sequence, we review some of the most important steps that constitute a standard protocol for the automatic land cover mapping from satellite images. Moreover, we delineate a methodological approach for the production and assessment of land cover maps from multitemporal satellite images that guides us in the production of a land cover map with high thematic accuracy for the study area. Finally, we develop a nonlinear harmonic model for fitting multispectral reflectances and vegetation indices time series from satellite images for numerous land cover classes. The simplified multitemporal information retrieved with the model proves adequate to describe the main land cover classes' characteristics and to predict the time evolution of land cover classes' individuals.

Keywords: Remote sensing, time series, multispectral images, land cover, separability analysis, classification, prediction, reference data, accuracy assessment.

LIST OF ABBREVIATIONS AND ACRONYMS

AG	Asymmetric Gaussian
AgRISTARS	Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing
ANN	Artificial Neural Network
ASAR	Advanced Synthetic Aperture Radar
AVHRR	Advanced Very High Resolution Radiometer
BMU	Best Matching Unit
BRDF	Bidirectional Reflectance Distribution Function
CLC	CORINE Land Cover
CORINE	Co-ordination of Information on the Environment
COS	*Carta de Ocupação do Solo* (Land Cover Map)
dB	Decibel
DGRF	*Direcção Geral dos Recursos Florestais* (Portuguese Forest Services)
DL	Double Logistic
EEA	European Environment Agency
ENS	*École Normale Supérieure*
ENVISAT	Environmental Satellite
EO	Earth Observation
ESA	European Space Agency
ETM	Enhanced Thematic Mapper
EU	European Union
EVI	Enhanced Vegetation Index
FAO	Food and Agriculture Organization
FDA	Fisher Discriminant Analysis
FR	Full Resolution
GIS	Geographic Information Systems
GLC	Global Land Cover
HRG	High Resolution Geometric
HT	Horvitz-Thomson
ICR	Interpretation Confidence Rating
IGP	*Instituto Geográfico Português* (Portuguese Geographic Institute)
INRIA	*Institut National de Recherche en Informatique et Automatique*
K-NN	K – Nearest Neighbor

LCCS	Land Cover Classification System
LCR	Location Confidence Rating
LDC	Linear Discriminant Classifier
LGN	Dutch land use database
LR	Low Resolution
LULC	Land Use and Land Cover
LVQ	Linear Vector Quantization
MARS	Monitoring Agriculture with Remote Sensing
MERIS	MEdium Resolution Imaging Spectrometer
MIR	Mid-Infrared
ML	Maximum Likelihood
MLP	Multi-Layer Perceptron
MMU	Minimum Mapping Unit
MODIS	Moderate Resolution Imaging Spectroradiometer
MRLC	Multi-Resolution Land Characteristics
MSE	Mean Square Error
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
NOAA	National Oceanic and Atmospheric Administration
NRMSE	Normalized Root Mean Square Error
PCA	Principal Component Analysis
Pdf	Probability density functions
RBF	Radially Basis Function
RMSE	Root Mean Square Error
RMSPE	Root Mean Square Prediction Error
RR	Reduced Resolution
RS	Remote Sensing
RSU	Remote Sensing Unit
SAR	Synthetic Aperture Radar
SD	Solar diffuser
SDSM	Solar diffuser stability monitor
SNR	Signal to Noise Ratio
SOM	Self-Organizing Map
SPOT	*Systeme Probatoire d'Observation de la Terre*
SVM	Support Vector Machine

SWIR	Shortwave Infrared
TM	Thematic Mapper
TOA	Top of the Atmosphere
US	United States
VI	Vegetation Index
VIS	Visible
WFAS	Wildland Fire Assessment System
WOY	Week Of the Year

TABLE OF CONTENTS

DEDICATION i
ACKNOWLEDGMENTS iii
RESUMO vii
ABSTRACT ix
LIST OF ABBREVIATIONS AND ACRONYMS xi
LIST OF TABLES xix
LIST OF FIGURES xxi
1 INTRODUCTION 1
1.1. Overall research objectives 1
1.2. Dissertation organization 1
1.3. Context 5
1.4. References 19
2 USE OF INTRA-ANNUAL SATELLITE IMAGERY TIME-SERIES FOR LAND COVER CHARACTERIZATION PURPOSE 25
Abstract 25
2.1. Introduction 25
2.2. Study area and data 27
2.3. Methodology 27
2.4. Results and discussion 29
2.5. Conclusions 36
2.6. Acknowledgements 36
2.7. References 36
3 CONTRIBUTION OF MULTISPECTRAL AND MULTITEMPORAL INFORMATION TO LAND COVER CLASSIFICATION FROM MODIS IMAGES 41
Abstract 41
3.1. Introduction 42
3.2. Study area and experimental data set 44
3.3. Data analysis 48
3.4. Results and discussion 55
3.5. Conclusion 63
3.6. Acknowledgements 63
3.7. References 64
4 SEPARABILITY ANALYSIS OF LAND COVER CLASSES AT REGIONAL SCALE: A COMPARATIVE STUDY OF MERIS AND MODIS DATA 71
Abstract 71

4.1. Introduction 71
4.2. Land cover nomenclature 72
4.3. Dataset 73
4.4. Methodology 76
4.5. Results and discussion 78
4.6. Conclusion 80
4.7. Acknowledgements 81
4.8. References 81
5 MERIS BASED LAND COVER CHARACTERIZATION: A COMPARATIVE STUDY 85
Abstract 85
5.1. Introduction 85
5.2. Study area and data set 88
5.3. Methodology 90
5.4. Results and discussion 96
5.5. Conclusion 99
5.6. Acknowledgements 100
5.7. References 100
6 MERIS BASED LAND COVER CLASSIFICATION WITH SELF-ORGANIZING MAPS: PRELIMINARY RESULTS 105
Abstract 105
6.1. Introduction 105
6.2. Study area and data set 107
6.3. Methodology 108
6.4. Results and discussion 111
6.5. Conclusion 113
6.6. Acknowledgements 114
6.7. References 114
7 SAMPLE DESIGN AND ANALYSIS FOR THEMATIC MAP ACCURACY ASSESSMENT: AN APPROACH BASED ON DOMAIN ESTIMATION FOR THE VALIDATION OF LAND COVER PRODUCTS 117
Abstract 117
7.1. Introduction 117
7.2. Survey sampling 118
7.3. Analysis and estimation 120
7.4. Acknowledgements 125
7.5. References 125

8 A REFERENCE SAMPLE DATABASE FOR THE ACCURACY ASSESSMENT OF MEDIUM SPATIAL RESOLUTION LAND COVER PRODUCTS IN PORTUGAL 127
Abstract 127
8.1. Introduction 127
8.2. Land Cover Nomenclature 128
8.3. Database Design 129
8.4. Database Analysis 133
8.5. Conclusion 134
8.6. Acknowledgments 134
8.7. References 134
9 MULTITEMPORAL MERIS IMAGES FOR LAND COVER MAPPING AT NATIONAL SCALE: THE CASE STUDY OF PORTUGAL 137
Abstract 137
9.1. Introduction 137
9.2. Land cover nomenclature 140
9.3. Study area 142
9.4. Dataset 142
9.5. Methodology 144
9.6. Results and Discussion 151
9.7. Conclusion 156
9.8. Acknowledgements 157
9.9. References 157
10 A NONLINEAR HARMONIC MODEL FOR FITTING SATELLITE IMAGES TIME SERIES: ANALYSIS AND PREDICTION OF LAND COVER DYNAMICS 163
Abstract 163
10.1. Introduction 164
10.2. Satellite images time series fitting 165
10.3. Case studies 172
10.4. Conclusion 185
10.5. Acknowledgment 187
10.6. References 187
11 CONCLUSIONS 191
11.1. Main outcomes 191
11.2. Main difficulties 196
11.3. Recommendations and future work 198

LIST OF TABLES

Table 1.1. MODIS, MERIS, NOAA-AVHRR and Landsat ETM+ main technical characteristics. 9

Table 2.1. User's and Producer's accuracies per class. This repartition corresponds to the SVM classification based on reflectances and vegetation indices obtained at a single date (August, 2000). 31

Table 2.2. Overall land cover classification accuracies obtained with combinatorial subsets of two dates 32

Table 2.3. Overall classification accuracies from the best single date classification, from the best two and three dates combinatorial sets and from the twelve dates association. The table shows the individual classification accuracy for each class (producer's accuracy) and overall accuracy. 34

Table 2.4. Z values to appraise the significance of the difference between classification accuracies from the best single date classification, from the best two and three dates combinatorial sets and from the twelve dates association. 35

Table 3.1. Land cover classes description, label and respective number of collected samples. 45

Table 3.2. Mahalanobis median distances $D(F_{p,q})$ between the nine land cover classes. $F_{p,q}$ corresponds to the feature subspace composed of the p channels $\{B_{i1},\ldots, B_{ip}\}$ measured at the q dates $\{t_{j1}, \ldots, t_{jq}\}$. 56

Table 3.3. Overall classification accuracy $C(F_{p,q})$ of the nine land cover classes, obtained with an optimized SVM trained and tested on $F_{p,q}$. Each $F_{p,q}$ corresponds to the feature subspace composed of the p channels $\{B_{i1},\ldots, B_{ip}\}$ measured at the q dates $\{t_{j1}, \ldots, t_{jq}\}$. 57

Table 3.4. User's and producer's accuracy corresponding to classes that mix easily: Natural Grassland (NG), Barren (B) and Shrubland (S). Values correspond to SVM classifications based on two different input feature spaces (see Table 3.2). 62

Table 4.1. LANDEO land cover nomenclature, class label and respective number of collected sample observations. 73

Table 4.2. Grouped classes with no average difference at 95% level of confidence. 79

Table 5.1. Land Cover nomenclature and classes description. 91

Table 5.2. Overall classification accuracies achieved with different classifiers and feature data sets. 98

Table 5.3. User's and producer's accuracy per class. This repartition corresponds to a SVM classification of FDA transformed data based on reflectances measured at a single date (August 14, 2004). 99

Table 6.1. Land cover classes and respective acronym. 108

Table 6.2. User's and producer's accuracy per class. This repartition corresponds to the classification of test samples performed with the SOM network with 1444 neurons. .. 113

Table 7.1. Sample error matrix. 122

Table 8.1. Land cover classes and respective number of collected sample observations. See Table 8.2 for columns titles. 129

Table 8.2. Attribute Information Collected for Interpreting Agreement Between Map and Reference Data. Adapted from [12]. 133

Table 8.3. ICR and LCR statistics for reference database..... 134

Table 9.1. LANDEO land cover nomenclature at the most disaggregated level..... 141

Table 9.2. CLC2000 level 3 nomenclature..... 143

Table 9.3. Reference sample sizes for training and testing. Test 1 and Test 2 represent, respectively, the number of observations per land cover class collected as primary and secondary reference. 145

Table 9.4. Percentage area occupied per mapped land cover class (N_h), and 95% confidence intervals for percentage area occupied per reference land cover class (N_g), overall accuracy (P), user's accuracy (P_h) and producer's accuracy (P_g). Absolute precision is represented by d. 151

Table 9.5. Estimated proportion ($\hat{N}_{hg}$) of matches between map and reference classes as a percentage of the total study area. 153

Table 9.6. Comparison between land cover map derived from 2005 multitemporal MERIS images and CLC2000. Each cell represents the classes' intersection (%) in relation to the total area of the territory. Intersections with correct agreement between the two nomenclatures are in bold face. 155

Table 10.1. Model parameters used to synthesize time series data. 169

Table 10.2. Mean SNR and SNR gain for model fits to input time series with different SNRs. 171

Table 10.3. Land cover classes, label and respective number of collected reference observations. 174

Table 10.4. *NRMSE*s (%) corresponding to the nonlinear, AG and DL functions' fits to spectral bands and VIs time series for each land cover class. 176

LIST OF FIGURES

Figure 2.1. Overall classification accuracy as a function of time. 30

Figure 2.2. Mean EVI values as function of time for “Natural grassland” (Ng), “Barren” (B), and “Rain Fed Herbaceous Crops” (RHC). See (33) for further details on mean profiles calculation..... 31

Figure 2.3. Overall classification accuracy per combinatorial subsets of two dates. Each box plot represents the lower quartile, median, and upper quartile of classification accuracy obtained for all pair-wise combinations of each month with each of the remaining eleven months images. Classification values that are considered as outliers are represented by “+”. 32

Figure 2.4. Overall classification accuracy obtained with observations from a single date and from combinatorial sets of two and three dates. For each month, only the best overall classification accuracies obtained by combining it with one and two other dates are displayed.. 34

Figure 3.1. Typical time series corresponding to the intra-annual evolution of reflectance in MOD09A1 spectral band B1 (red wavelength) for a pixel area of the Rainfed Crop (RC) class. Crosses correspond to the raw observed data. Dots show the output of a time-varying windowed median filter (we used a gaussian window with a width corresponding to a six weeks duration). 49

Figure 3.2. EVI time profiles obtained by averaging the sample belonging to the same land cover class. The three displayed classes, Natural Grassland (NG), Barren (B) and Rainfed Crops (RC), mix up in summer time, whereas they are perfectly distinguishable in winter.. 57

Figure 3.3. Overall classification accuracy progression when increasing the number of measurements along time. For the two situations: comprising and discarding the vegetation indices, classifications based on a single (best) channel and based on all channels are presented. 61

Figure 5.1. Spatial distribution of collected samples. 92

Figure 5.2. Mean spectral signatures for land cover classes derived from the MERIS image of August 14th, 2004. 97

Figure 9.1. Land cover map for Continental Portugal derived from multitemporal MERIS images of 2005. 152

Figure 10.1. Average fitting error (%), $< e_N >_\theta$, corresponding to the same N. 170

Figure 10.2. Nonlinear model fits for time series synthesized with different SNRs..... 171

Figure 10.3. Observed EVI time series for three Rice crops (23) individuals and the respective nonlinear, AG and DL fits. 178

Figure 10.4. Observed EVI time series for three Needleleaved forest (321) individuals and the respective nonlinear, AG and DL fits. 178

Figure 10.5. Observed EVI time series for two Eucalyptus forest (3112) individuals and the respective nonlinear, AG and DL fits. 179

Figure 10.6. Rice crops (23) phenological attributes. a) NDVI time series and respective model fit for a single observation; b) first and second derivatives of modelled time series as function of time. Vertical bars represent the times for: Transplanting t^T; Onset of maturity t^{OnM}; Maximum greenness t^M; Offset of maturity t^{OfM}; Harvesting t^H. 180

Figure 10.7. Estimated Rice crops (23) sample probability distributions for: a) Transplanting time t^T; b) Maturity time t^M; c) Harvesting time t^H. Vertical bars represent the limits for the phenological reference key-dates provided in [36] for rice crops in Portugal. 181

Figure 10.8. Rice crops (23) sample *RMSPE*s for: a) dates of maturity t^M as function of t_n; b) dates of harvesting t^H as function of t_n. 183

Figure 10.9. Estimated Rice crops (23) sample probability distribution for eN_n^M, where $t_n = t^M - 6$. 184

Figure 10.10. Estimated Shrubland (34) sample probability distributions for: a) eT_{20}^m; b) eN_{20}^m. 184

Figure 10.11. *NDVI* time series model fits for two Shrubland (34) individuals and respective model predictions for t_{20}. The rectangles represent the standard deviations for eT_{20}^m and eN_{20}^m. 185

1 INTRODUCTION

1.1. Overall research objectives

The broad objective of this research is to investigate the usefulness of intra-annual time series sets of multispectral satellite images with medium spatial resolution for the automatic assessment of land cover distribution and dynamics at the national scale. The goal is to give response to some of the land cover characterization tasks that were attributed to the Portuguese Geographic Institute (IGP) in the framework of the Portuguese R&D National Strategy for Space (approved by the Science Ministry on 5th March 2004), namely the planning of strategies for the regular mapping of land cover distribution and analysis of land cover dynamics at the national scale.

In a series of nine companion studies, the work described here evaluates the suitability of high temporal and spectral resolution satellite images acquired by the Moderate Resolution Imaging Spectroradiometer (MODIS) and MEdium Resolution Imaging Spectrometer (MERIS) for the periodic characterization of land cover in Continental Portugal. In general, we review some of the major problems related with the numerous steps that constitute a standard protocol for the automatic land cover mapping from satellite images, namely: selection of the most adequate set of Earth Observation (EO) data, nomenclature definition and respective sample data collection, automatic classification problem, and accuracy assessment process. In detail, we aim at understanding and appraising the value of multi-temporality in the context of automatic land cover classification from multispectral satellite images. We first focus on the analysis of the contribution of multitemporal and multispectral information from satellite images for land cover classes' discrimination. In the sequence, we investigate a mathematical function to methodically describe the intra-annual evolution of multispectral reflectances and vegetation indices (VI) time series sets from satellite images for different land cover classes. The goal is to optimize the contribution of high frequency multitemporal information from multispectral satellite images for the characterization and prediction of land cover classes distribution and dynamics.

1.2. Dissertation organization

This dissertation is composed of nine stand-alone manuscripts published in Peer-Reviewed Journals and International Conferences and Workshops within the Remote Sensing (RS) and Geographic Information Systems (GIS) scientific communities. The manuscripts are introduced by this opening Chapter and followed by a Chapter of conclusions, limitations and recommendations. Slight formatting differences in each Chapter reflect the different

requirements of the respective Peer-Reviewed Journals (4) or Conferences (5) to which the stand-alone manuscripts have been submitted for publication. The individual manuscripts were ordered in this dissertation so as to reproduce the line of research in a natural time sequence, i.e. such that the outcomes of each stand-alone publication directly link with the research premises of the following manuscripts. The order of the publications in this dissertation and their respective citations are:

- **Carrão**, H., Gonçalves, P. and Caetano, M. (2007), Use of intra-annual satellite imagery time-series for land cover characterization purpose. *EARSeL eProceedings*, 6, 1-11.
- **Carrão**, H., Gonçalves, P. and Caetano, M. (2008), Contribution of multispectral and multitemporal information from MODIS images to land cover classification. *Remote Sensing of Environment*, 112, 986-997.
- **Carrão**, H., Sarmento, P., Araújo, A. and Caetano, M. (2007), Separability analysis of land cover classes at regional scale: a comparative study of MERIS and MODIS data. In *Proceedings of the ENVISAT Symposium* (unpaginated CD-ROM), 23-27 April 2007, Montreaux, Switzerland.
- **Carrão**, H., Gonçalves, P. and Caetano, M. (2006), MERIS Based Land Cover Characterization: A Comparative Study. In *Proceedings of the ASPRS 2006 Annual Conference - Prospecting for Geospatial Information Integration* (unpaginated CD-ROM), 1-6 May 2006, Reno, Nevada, USA.
- **Carrão**, H., Capão, L., Bação, F. and Caetano, M. (2006), MERIS Based Land Cover Classification With Self-Organizing Maps: Preliminary Results. In *Proceedings of the 2nd EARSeL SIG Workshop on Land Use & Land Cover* (unpaginated CD-ROM), 28-30 September 2006, Bonn, Germany.
- **Carrão**, H., Caetano, M. and Coelho, P. (2007), Sample Design and Analysis for Thematic Map Accuracy Assessment: An Approach Based on Domain Estimation for the Validation of Land Cover Products. In *Proceedings of the 32nd International Symposium on Remote Sensing of Environment* (unpaginated CD-ROM), 25-29 June 2007, San Jose, Costa Rica.
- **Carrão**, H., Araújo, A., Cerdeira, C., Sarmento, P., Capão, L. and Caetano, M. (2007), A reference sample database for the accuracy assessment of medium spatial resolution land cover products in Portugal. *IEEE International Geoscience and*

Remote Sensing Symposium (IGARSS´2007), 23-28 July 2007, Barcelona, Spain, 3967-3970.

- **Carrão**, H., Araújo, A., Gonçalves, P. and Caetano, M. (2009), Multitemporal MERIS images for land cover mapping at national scale: the case study of Portugal. *International Journal of Remote Sensing*, to be published.
- **Carrão**, H., Gonçalves, P. and Caetano, M. (2009), A nonlinear harmonic model for fitting satellite images time series: analysis and prediction of land cover dynamics. *IEEE Transactions on Geoscience and Remote Sensing*, submitted.

Furthermore, in the framework of this research, several parallel studies were developed in close collaboration with master students at the Remote Sensing Unit (RSU) of the Portuguese Geographic Institute (IGP) and were published as manuscripts in Peer-Reviewed Journals (1) and International Conferences (16) within the Remote Sensing community. Next, we present the complete list of published manuscripts:

- Sarmento, P., **Carrão**, H., Caetano, M. and Stehman, S. V. (2009), Incorporating reference classification uncertainty into the analysis of land cover accuracy. *International Journal of Remote Sensing*, to be published.
- Araújo, A., **Carrão**, H. and Caetano, M. (2009), An operational approach for annual land cover mapping at the national scale with MERIS images. In D. Maktav (Ed.), *Remote Sensing for a Changing Europe* (pp. 602-609). Amsterdam: IOS Press.
- Costa, H., **Carrão**, H., Bação, F. and Caetano, M. (2008), Land cover classification in Portugal with multitemporal AWiFS images: a comparative study. In D. Maktav (Ed.), *Remote Sensing for a Changing Europe* (pp. 239-246). Amsterdam: IOS Press.
- Araújo, A., **Carrão**, H. and Caetano, M. (2008), Using estimated land cover classes accuracies in multi-temporal classification. In M. Gomarasca (Ed.), *Geoinformation in Europe* (pp. 255-365). Rotterdam: Millpress.
- Sarmento, P., **Carrão**, H. and Caetano, M. (2008), A fuzzy synthetic evaluation approach for land cover cartography accuracy assessment. In J. Zang and M. Goodchild (Ed.), *Spatial Uncertainty* (pp. 348-355). Liverpool: World Academic Union.
- **Carrão**, H., Araújo, A. and Caetano, M. (2008), Land cover classification in Portugal with intra-annual time series of MERIS images. In *Proceedings of the 2nd*

MERIS/(A)ATSR User Workshop, 22-26 September 2008, ESA/ESRIN Frascati (Rome), Italy.

- Costa, H., Araújo, A., **Carrão**, H. and Caetano, M. (2008), Influência das características técnicas das imagens de satélite na produção de cartografia de ocupação do solo: estudo baseado em imagens MERIS e AwiFS. In *Actas do X Encontro de Utilizadores de Informação Geográfica* (eSIG'08), 14-16 May 2008, Oeiras, Portugal, in press.
- Sarmento, P., **Carrão**, H. and Caetano, M. (2008), Avaliação da exactidão temática de cartografias de ocupação do solo através de funções fuzzy: primeira abordagem. In *Actas do X Encontro de Utilizadores de Informação Geográfica* (eSIG'08), 14-16 May 2008, Oeiras, Portugal, in press.
- Capão, L., **Carrão**, H., Araújo, A. and Caetano, M. (2007), An approach for land cover mapping with multi-temporal MERIS imagery. *IEEE International Geoscience and Remote Sensing Symposium* (IGARSS´2007), 23-28 July 2007, Barcelona, Spain, 3836-3839.
- **Carrão**, H., Sarmento, P. Araújo, A. and Caetano, M. (2007), Retrieving land cover information from MERIS and MODIS data: a comparative study for landscape characterization in Portugal. *IEEE International Geoscience and Remote Sensing Symposium* (IGARSS´2007), 23-28 July 2007, Barcelona, Spain, 1271-1274.
- **Carrão**, H., Gonçalves, P. and Caetano, M. (2007), Use of intra-annual satellite imagery time-series for land cover characterization purpose. In Z. Bocheneck (Ed.), *New Developments and Challenges in Remote Sensing* (pp. 355-365). Rotterdam: Millpress.
- **Carrão**, H. and Caetano, M. (2007), A comparison of MERIS and MODIS data for land cover characterization. In *Anais XIII Simpósio Brasileiro de Sensoriamento Remoto*, Florianópolis, Brasil, 21-26 Abril 2007, INPE, p. 5631-5633.
- Gonçalves, P., **Carrão**, H. and Caetano, M. (2007), Parametric model for intra-annual reflectance time series. In Z. Bocheneck (Ed.), *New Developments and Challenges in Remote Sensing* (pp. 367-375). Rotterdam: Millpress.
- Pinheiro, A., **Carrão**, H. and Caetano, M. (2007), Evaluation of ASAR And Optical Data Synergy For High Resolution Land Cover Mapping In Portugal. *IEEE International Geoscience and Remote Sensing Symposium* (IGARSS´2007), 23-28 July 2007, Barcelona, Spain, 1517-1520.

- Pinheiro, A., **Carrão**, H., Gonçalves, P. and Caetano, M. (2007), A Toolbox for multi-temporal analysis of satellite imagery, In Z. Bocheneck (Ed.), *New Developments and Challenges in Remote Sensing* (pp. 395-404). Rotterdam: Millpress.
- Cerdeira, C., Araújo, A., **Carrão**, H. and Caetano, M. (2006), Validação das Cartografias Globais de Ocupação do Solo, GLC2000 e MOD12Q1, para Portugal Continental. In *Actas do IX Encontro de Utilizadores de Informação Geográfica* (unpaginated CD-ROM), 15-17 November 2006, Oeiras, Portugal.
- Gonçalves, P., **Carrão**, H., Pinheiro, A. and Caetano, M. (2006), Land cover classification with Support Vector Machine Applied to MODIS imagery. In A. Marçal (Ed.), *Global Developments in Environmental Earth Observation from Space* (pp. 517-526). Rotterdam: Millpress.

1.3. Context

1.3.1. Motivation

Land cover, i.e. "the observed biophysical cover on the Earth's surface" (Di Gregorio and Jansen 2000), is key environmental information. It embraces, for example, the quantity and type of surface vegetation, water, and earth materials (Meyer and Turner 1994). Land cover is generally a reflection of the local climate and landforms, though it can be altered by the human use of land. It is important for many scientific domains, resource management and policy purposes, and for a range of human activities (Comber 2008).

Land cover varies at a range of spatial scales, from local to global, and at temporal frequencies of days to millennia (Cihlar 2000). Variations in the composition and distribution of land cover types play an active role in the Earth’s surface energy and water balance, as well as in the carbon cycle (Penner 1994, Kalnay and Cai 2003). In turn, land cover changes can arise in response to natural hazards (drought, wind, floods, and rain erosion) and anthropic stress (industry, overgrazing, fires, land abandonment) (Veldkamp and Lambin 2001).

The periodic assessment of land cover distribution and of land cover dynamics is considered as the most important variable for studying the time evolution of many biophysical processes occurring in the Earth (Vitousek 1994). Land cover dynamics feed back on climate changes, making it possible to model the interactions between human activities, the Earth’s surface elements, and different natural and human-induced ecosystems (Foody 2002). In particular, the analysis of land cover dynamics is a break opportunity for capturing environmental

pressures, both natural and anthropogenic, that lean towards many physical processes and diffusive phenomena over different time periods and spatial scales (Foley *et al*. 2005).

As the need for environmental understanding, management and planning became urgent to mitigate several negative human impacts on the world's landscapes, an accompanying call for updated land cover information emerged in parallel (Loveland *et al.* 2000, Rogan and Chen 2004). Under this framework, land cover mapping has an important role as it can be a very useful tool for studying the dynamic properties of the Earth's environment (Stehman *et al.* 2000). Land cover maps provide important and detailed information about the Earth's surface that is commonly used for ecological monitoring, natural resources assessment, and policies development (Foody 2002). Land cover maps are easily understandable, managed and combined with other types of information (Veldkamp and Lambin 2001). Therefore, modelling the land cover dynamics, especially if done in a spatially explicit, integrated and multi-scale manner, is an important technique for the projection of alternative pathways into the future, for conducting experiments that test our understanding of environmental key processes, and for describing the latter in quantitative terms (Lambin *et al.* 2000).

On the one hand, maps of the distribution and status of the Earth's surface elements are nowadays critical for complex parameterization of climate, agricultural and epidemiological models, as well as for simple local to global scale forestry, industrial, urban and social applications (Muchoney *et al.* 2000). On the other hand, the ability to inventory and map land cover dynamics and to monitor land cover changes is required for, among other things, modeling the carbon and hydrologic cycles, studying land surface-climate interactions, and establishing rates of deforestation (Foley *et al*. 2005). Indeed, mapping land cover time evolution, namely in forests, grasslands, wetlands and agricultural lands, contributes significantly to the study of inter-annual land cover dynamics which have a direct impact on the carbon source and sink estimations and predictions (Clevers *et al.* 2007). Therefore, to protect the environment and to ensure sustainable use of natural resources, a wide variety of national and international mechanisms have been established for environmental monitoring, which on their turn have resulted on many land cover mapping activities. Examples are the European Union (EU) Habitats Directive (Wils 1994), the EU Common Agricultural Policy (Olesen and Bindi 2002) and the Kyoto Protocol (Steffen *et al.* 1998).

The specific environmental variables that are connected with land cover vary by application, discipline and model. Global land cover maps are adequate for the most part of global change research objectives, including the assessment of current global ecological conditions and the simulation of future environmental scenarios that ultimately lead to public policy

development (Loveland *et al*. 2000). However, at global scales, we can only expect to successfully monitor the status of pre-established impacts on the environment, but not to detect and effectively plan the mitigation of the causing disturbances at the appropriate times (Veldkamp and Lambin 2001). In contrast, as the spatial resolution of land cover maps becomes finer, the direct actors of land cover change can be identified and process-based relationships can be determined. (Hansen *et al*. 2000, Veldkamp and Lambin 2001). Therefore, for the establishment of effective environmental protection programs and land cover management applications, there is a need for land cover maps produced at local to regional scales (Skole and Cochrane 2004). Land cover dynamics that are on the basis of the environmental global stresses have important spatial attributes that appear only at those finer scales (Foley *et al*. 2005). Indeed, relevant land cover changes induced by human activities occur at typical spatial scales in the order of 250-500 m (Zhan *et al.* 2002). Thus, the development of dynamic models of land cover change for the establishment of national policies for, e.g. forest fire prevention, agricultural practices, water consumption, coastal protection and drought assessment, involves a closer examination of national scale processes (Mücher *et al.* 2000).

Independently of the spatial scale of analysis, it is extremely difficult to generate spatially and temporally coherent land cover products for monitoring and understanding land surface events from time irregular and geographically sparse records collected at ground stations (Anderson 1976). Because of the synoptic geographical coverage and frequent temporal acquisition of satellite images' observations, remotely-sensed data possess significant potential for automatic mapping of land cover distribution and dynamics from local to global scales (Rogan and Chen 2004). Satellite images provide more economical snapshots of the Earth's surface which can be acquired over continuous wide spatial areas. Maps produced from the automatic classification of multispectral satellite images data have the advantages of frequent acquisition and internal consistency of remotely-sensed data sets over traditional ground-based methods (Stehman *et al.* 2000, Foody 2002, Moran *et al.* 2004). Indeed, a primary reason for attempting to create maps from remotely-sensed data sets is the potential for the regular production of more accurate and consistent land cover information than past efforts compiled from ground-based methods (Loveland *et al*. 2000).

Time series of multispectral remote sensing data with high frequency of acquisition allow the prediction of land cover changes timing, opening new opportunities to better link macro-economic perspectives with land cover distribution and dynamics, and to better understand time delays in land cover responses to socio-economic or natural perturbations (Kaufmann and Seto 2001). However, automatic methods for the characterization of land cover

distribution and dynamics have, to date, made only limited use of multitemporal remotely sensed image data (Homer *et al.* 2004). Studies reporting the use of time series sets of satellite images for land cover classification often include relatively few dates, possibly due to a lack of cloud-free images, cost and processing requirements of existing remotely sensed data sets (Knight *et al.* 2006).

New satellite data

With the launch of MODIS and MERIS sensors, a broad range of new opportunities for automatic land cover mapping at the national scale started (Rast *et al.* 1999, Zhan *et al.* 2002, Rogan and Chen 2004). Although these data are too coarse for the purposes of local level land planning and management, they serve as a general product for national agencies (Rogan and Chen 2004). These images exhibit enhanced spectral and temporal resolutions, wide geographical coverage and improved atmospheric correction (Borak and Strahler 1999, Curran and Steele, 2005). Therefore, the gap between fine and coarse spatial resolution sensors can now be filled with imagery at MODIS and MERIS spatial resolutions, i.e. 500 and 300 m, respectively (Clevers *et al.* 2007). Indeed, both sensors overcome the spatial and spectral resolutions of NOAA-AVHRR sensor, and the temporal resolution of Landsat ETM+ sensor, thus promoting the regular production of land cover maps at the national scale through automatic classification of satellite images (Borak and Strahler 1999, Clevers *et al.* 2007). In Table 1.1, one presents the main technical characteristics of MODIS, MERIS, NOAA-AVHRR and Landsat ETM+ images.

On the one hand, the spectral and spatial resolutions of MERIS and MODIS images overcome those of NOAA-AVHRR and allow for distinguishing between more similar land cover classes that characterize the landscape features at the national scale (Borak and Strahler 1999, Verstraete *et al.* 1999). MODIS and MERIS images have approximately the same temporal resolution as those acquired by NOAA-AVHRR sensor, but higher spectral and spatial resolutions (Table 1.1). Thus, the thematic content of land cover maps nomenclatures can now be more comprehensive than it was in the past with maps derived from NOAA-AVHRR; the produced land cover maps have improved geometric resolutions that match the landscape characteristics at finer spatial scales. These technical conditions suggest that it is now possible to develop new automatic classification approaches based on the higher spectral and spatial resolutions of MODIS and MERIS images, which will in turn increase the discrimination between a larger set of land cover classes of interest (Borack and Strahler 1999, Clevers *et al.* 2007).

Table 1.1. MODIS, MERIS, NOAA-AVHRR and Landsat ETM+ main technical characteristics.

	Sensor			
	MODIS	**MERIS**	**NOAA-AVHRR**	**Landsat ETM+**
Spatial resolution	250 m (spectral bands 1-2) 500 m (spectral bands 3-7) 1000 m (spectral bands 8-36)	300 m (spectral bands 1-15) 1200 m (spectral bands 1-15)	1100 m (spectral bands 1-5)	30 m (spectral bands 1-5, 7) 60 m (spectral band 6)
Temporal resolution	2 days	3 days	Twice-daily	16 days
Spectral resolution	405-14.385 nm (36 spectral bands) Visible (10 spectral bands) Reflective Infra-Red (9 spectral bands) Thermal Infra-Red (17 spectral bands)	390-1040 nm (15 spectral bands) Visible (9 spectral bands) Reflective Infra-Red (6 spectral bands)	580-12.500 nm (5 spectral bands) Visible (1 spectral band) Reflective Infra-Red (1 spectral bands) Thermal Infra-Red (3 spectral bands)	450-12.500 nm (7 spectral bands) Visible (3 spectral band) Reflective Infra-Red (3 spectral bands) Thermal Infra-Red (1 spectral band)
Radiometric resolution	12 bits	12 bits	10 bits	8 bits

On the other hand, in comparison with Landsat ETM+ sensor, the enhanced temporal characteristics of MODIS and MERIS sensors stimulate the development of innovative automated land cover classification approaches that take advantage of standard time series analyses' techniques and/or combine the high temporal and spectral features of the respective remotely-sensed data sets (Zhang *et al.* 2003, Hermance *et al.* 2007). Indeed, multispectral time series sets of medium spatial resolution satellite images can be very useful for distinguishing between land cover classes that are spectrally similar at times (Borak and Strahler 1999, Maxwell *et al.* 2002). However, only a limited number of studies to date have explored the potential of employing a complete year of uninterrupted satellite images as the basis for land cover characterization (Knight *et al.* 2006). In the past, these approaches were not projected because it was extremely difficult to collect regularly time spaced and consistent intra-annual time series data sets of multispectral satellite images with adequate spatial resolution and covering the whole national territories (Moulin *et al.* 1998). The reason was that due to cloud contamination, many of the sparsely collected images were discarded from the time series because of the atmospheric noise corrupting the data (e.g. Landsat ETM+). Therefore, time series sets of satellite images were essentially noise contributing to poor classification accuracy (Knight *et al.* 2006).

1.3.2. Problem statement

Whereas the value of satellite images for frequent automatic mapping of land cover distribution and assessment of land cover dynamics is evident, and the need for updated land cover maps is clear, there is a lack of land cover products for national land cover applications. To effectively track land cover composition over national territories and monitor effective land cover changes, it is important that the technical characteristics of land cover products, namely the thematic content, the spatial scale, the date and periodicity of production, and the respective thematic accuracy are adequate (Foody 2002, Pal and Mather 2003, Treitz and Rogan 2004). Indeed, an important concern about land cover mapping for numerous scientific, policy, planning and management purposes is the consistency and reproducibility of the derived land cover products (Cihlar 2000). If the technical characteristics of land cover maps vary across time and space, then it is impossible to properly monitor land cover changes. Therefore, the active involvement of local experts can also help to produce an accurate and repeatable land cover characterization database (Giri *et al.* 2005).

Current land cover characterization programs

In the last years, several operational land cover mapping programs were established at global (e.g. MODIS Land Cover (MOD12Q1; Hodges 2001, Friedl *et al.* 2002), Global Land Cover 2000 (GLC2000; Bartholomé and Belward 2005)), regional (e.g. Co-ordination of Information on the Environment (CORINE) Land Cover map for 2000 (CLC2000; EEA 2002), National Land Cover Database (NLCD 2001; Homer *et al.* 2004)) and national (e.g. *Carta de Ocupação do Solo* 1990 (COS'90, IGP 2009)) scales to provide users with more accurate, consistent and updated free of charge land cover data sets. These initiatives, funded by both national and international institutions, focused on mapping land cover distribution and detecting land cover changes at different spatial and temporal resolutions. However, there exist two main issues: 1) technical characteristics at whole land cover maps are not defined for national scale applications; and 2) most appropriate land cover maps are not frequently updated. Still, the lack of more suitable products favors the frequent use of those free of charge land cover maps by local users within national applications. For that reason, it is not surprising that the outcomes of geographical applications for particular areas of interest are at times biased (Giri *et al.* 2005), and the subsequent political decisions are inconsistent with the real world conditions. There are numerous technical issues that hamper the effective use of those data sets as input features for national applications on a regular basis and that we wish to review with this study. Next, we introduce some of the main drawbacks of those land cover products:

(1) *Periodicity of production*. Land cover information produced in the framework of the existing mapping programs is not compatible with the temporal resolution of land cover dynamics. To serve as an input feature for environmental models and resource applications, the land cover information should be updated at least on a yearly basis (e.g. for forest fire prevention (Caetano *et al.* 2004)). However, there are several factors that hinder the regular production of land cover maps from satellite images, namely those related with economic and time costs of the used mapping protocols.

On the one hand, and regarding the land cover maps produced by visual interpretation of satellite images, such as CLC2000 and COS'90, their regular production is not practicable on an operational basis because of the time used for: 1) image interpreters training; 2) manual delineation of land cover polygons on computer screen; and 3) harmonization of the final product to avoid discrepancies between individual image interpretations. Moreover, because the production process is labour-intensive, it is also economically expensive. So, the manual production of land cover databases from visual interpretation of satellite images no longer provides the map requirements on due time for most national agencies.

On the other hand, and looking now at the land cover maps produced by automatic classification of satellite images (e.g. MOD12Q1, GLC2000, and NLCD 2001), there is a difficulty in their regular production that is mainly related to the respective scale of application. Because these products are derived at global and regional scales, there are several specific steps within the standard mapping protocol that make their production extremely expensive, slow and difficult to support on a regular basis, namely: 1) the worldwide training sample acquisition; 2) the image compositing process; 3) the geometric, atmospheric and radiometric corrections (as well as the radiometric harmonization of many satellite images); 4) the classification process (usually, the characterization of land cover areas over the global surface is performed through multi-step approaches, e.g. Bartholomé and Belward (2005)); and 5) the final harmonization of the global product (i.e. post-processing).

The thematic detail of the adopted land cover nomenclature for both manual and automatically derived products has also a strong consequence on the frequency of land cover maps production: detailed thematic contents make the maps production slower than the production of maps with broad nomenclatures.

(2) *Thematic content*. The land cover information may be needed for few cover types, for all land cover types and at the same levels of detail, or tailored for specific

models requirements (Cihlar 2000). From a national application perspective, we remark that land cover nomenclatures of global and even regional products may not be adequate to use as input features for most tasks. The thematic content of global maps is extremely generic to assemble in few cover types all regional specificities occurring on the Earth's surface. Therefore, the broad cover classes of global maps' nomenclatures are most of the times useless to depict national specific land cover types (Giri *et al.* 2005). As a consequence, the thematic information in global land cover products is insufficient to detect the land cover changes and properly follow the dynamics of human induced disturbances that occur at finer scales (Rogan and Chen 2004).

Moreover, the nomenclatures of the existing land cover products are difficult to link (Giri *et al.* 2005). The national, regional and global maps are produced by different institutions and to address different and specific requirements. As a consequence, and even if produced at sequential times, the different maps may not be used to monitor land cover dynamics because their thematic content is different (Giri *et al.* 2005). This problem emphasizes the need for a worldwide hierarchical land cover nomenclature that comprises all global cover types at the base level and in turn that splits itself naturally into specific cover types at different superior levels according to the geographic region of interest and spatial scale of analysis.

(3) *Spatial resolution and geographical coverage*. Thematic land cover maps derived from satellite images are usually distributed on a regular spaced lattice that corresponds to the satellite sensor spatial resolution for images acquisition. Accordingly, the spatial resolution of land cover maps derived from coarse spatial resolution satellite images, i.e. global maps, is not adequate to represent most of landscape characteristics and track land cover changes at the national and regional scales (Treitz and Rogan 2004). Indeed, in a previous study (Carrão 2002), we showed that the landscape features in Continental Portugal can be perfectly represented with minimum mapping units smaller than 9 ha, i.e. with maps that have spatial resolutions higher than 300 m. Land cover maps with $1km^2$, such as MOD12Q1 and GLC2000, are thus not adequate to depict most of the land cover characteristics and changes that occur at the national scales. Satellite images with low spatial resolutions limit the production of land cover maps for monitoring purposes at the national scale due to the finer resolution at which most land cover changes take place (Mücher *et al.* 2000).

To overcome this situation, many scientific studies at the regional and national geographic scales have been projected with satellite images of 10-30m spatial resolutions (Cohen and Goward 2004). However, the use of these data is not appropriate in the context of an operational mapping program for national applications due to their limited spatial extend and low revisit time (Clevers *et al.* 2007). Because high spatial resolution satellite images (e.g. Landsat ETM+) just cover a small proportion of the national territories, it is also necessary to merge a set of neighbor images acquired under different conditions to completely cover the national surface. So, as radiometric characteristics of different images are difficult to harmonize, the existing differences can affect the classification results (Cihlar 2000). The classifiers performances are usually diminished because pixels of the same land cover class are affected by different radiometric disturbances in different images and present dissimilar spectral characteristics (Lu and Weng 2007).

(4) *Thematic accuracy*. Remote sensing data has a considerable potential for the regular provision of land cover maps, but the accuracy of the maps derived is often viewed as insufficient. It is reckoned that global land cover products present smaller overall accuracies for certain particular areas of interest than depicted by producers. For example, MOD12Q1 producers claim overall accuracies between 75-80% (Giri *et al*. 2005), whereas the validation of this product for Portugal does not exceed 57% accuracy (Cerdeira *et al*. 2006). In addition, comparison of MOD12Q1 with CLC2000, within the same region, showed an only 50% agreement between the two products (Caetano and Araújo 2006). Whereas the required thematic accuracy for different applications may be different, it is obvious that 50% overall accuracy is far from what is required for the evaluation of stresses on the environment and establishment of national policies concerning the prevention and mitigation of natural and human-induced pressures.

The low overall and specific thematic accuracies attained for these global land cover maps at localized geographical regions is due to several methodological issues and intrinsic remotely-sensed data factors, which provoke individual and/or combinatorial problems on the production and validation processes of map products (Giri *et al*. 2005). On the one hand, to properly distinguish between characteristic land cover classes through automatic classification of satellite images, physical measurements of Earth's surface collected by sensors must be adequate. It is extremely difficult, if not impossible, to automatically distinguish between the individuals of numerous land cover classes if the information contained in the

remotely-sensed data sets, which serve as input features for the classifier, is not sufficient to detect their subtle differences (Borak and Strahler 1999). On the other hand, but regarding also the thematic content of land cover nomenclatures, we can not afford to discriminate between neighbor land cover classes' individuals that are smaller than the spatial resolution of satellite images used for classification (Treitz and Rogan 2004).

Furthermore, there are also problems related to the selected classification approaches, which are not adequate to discriminate all classes in the land cover nomenclature (Foody 2002, Pal and Mather 2003).

Finally, the accuracy assessment protocols used to evaluate the thematic quality of land cover maps are usually not adequate. Users and producers sometimes resort to accuracy assessment approaches that do not respect the probabilistic designs that were used to collect the testing samples, whereas a correctly performed statistical accuracy assessment is the core of the validation process (Stehman and Czaplewski 1998). The most common errors for reporting maps correctness concern the incorrect calculation of the variance for the accuracy estimators (Stehman 1996). The complications related with the correct variance estimation are mainly due to the comparison of the map classification with (Stehman 2001, Foody 2002): 1) the classifier training sample; 2) the testing samples deterministically selected over the study area and that are geographically correlated; 3) other existing land cover maps that are also representations of the truth and not the reference data. Moreover, and most of the times, estimators properties are not correctly used for the calculation of the respective confidence intervals (Stehman and Czaplewski 1998).

In Chapter 1.3.1, we claimed that the technical characteristics of MODIS and MERIS sensors are adequate to tackle most of the presented land cover characterization issues. Therefore, in this research, we focus on the analysis of multispectral time series sets of MODIS and MERIS images for the regular production of land cover products for national scale applications and on the use of probabilistic techniques for the statistical accuracy assessment of those products. Nevertheless, we need to overcome some data analysis problems that still hamper their effective use and that we introduce afterwards.

Data analysis

The technical characteristics of MODIS and MERIS images, namely their spatial, temporal and spectral resolutions, suggest that the automatic production of land cover maps for national scale applications is now feasible on a regular basis. Moreover, the monitoring of

land cover dynamics can be better guaranteed through the time series analysis of images acquired by those sensors (Clevers *et al.* 2007). Nevertheless, we can not simply take a time series set of daily acquired medium spatial resolution satellite images with high spectral resolution, and expect that a standard classifier will accurately distinguish between the land cover classes of interest! There is no supervised or unsupervised black box like this! Thus, there is a growing necessity for the development of methods to extract the information contained in those remotely-sensed data sets in the framework of real world operational applications.

Because the dimensional feature spaces of multispectral time series sets of MODIS and MERIS remotely-sensed data are extremely large, the use of standard classifiers to automatically discriminate between individuals of numerous land cover types is not straightforward (Lu and Weng 2007). In fact, the use of multitemporal data from medium spatial resolution satellite images increases the dimensional feature space of multispectral input data sets for classification, while increasing the classification complexity (Maxwell *et al.* 2002). Therefore, if the standard parametric land cover classification approaches were unsuccessful before for smaller data sets, it is likely that they still fail for classifying these new high-dimensional remotely-sensed data sets. The main problem is that the size of the training data set, which is required to manage the within-classes' variances in a standard supervised parametric automatic classification process, increases with the augment of both multitemporal and multispectral dimensions of the medium spatial resolution satellite images (Hughes 1968, Jackson and Landgrebe 2001, Guyon and Elissee 2003). Because training sample acquisition is time-consuming, labour-intensive and economically expensive, the collection process is usually disregarded by producers and the sample size is not enough to deal with the classification task (Mathur and Foody 2006). Therefore, standard algorithms often fail to deliver high classification accuracies and tend to suffer from the problem of the "curse of dimensionality" (Hughes 1968).

Nowadays, there are already innovative classification techniques, such as the Support Vector Machines (SVM) (Vapnik 1998), that theoretically overcome the large dimensional feature space problem associated with the standard automatic classification techniques (Müller *et al.* 2001). To be precise, these algorithms require a small number of training individuals to work properly under large dimensional feature spaces (Foody and Mathur 2004). However, these learning machines are not user friendly because they rely on several heuristic parameters that need to be tuned for each specific application and image datasets (Meyer *et al*. 2003). In addition, as their processing time is extraordinary huge for testing scenarios performed on restrict geographical areas, their application on an operational basis for national tasks may

take processing times that are not yet affordable (Huang *et al.* 2002). Therefore, although they work very well under experimental conditions (Meyer *et al.* 2003), previous studies suggest that their effective use on an operational remote sensing application is far from optimal (e.g. Huang *et al.* 2002).

Nonetheless, if we do not collect the appropriate set of observations for training the SVM classifier, within the context of a large dimensional feature data set, it is also difficult, if not impossible, to go from experimental to real scenarios without generalization lost (Pal and Mather 2003). The problem is that we may incur into training overfitting (Van Niel *et al.* 2005); the training accuracy is perfect, but the learning machine lacks generalization to deal with multiple real world within-classes' differences (Landgrebe 2002).

Therefore, the answer for the problem related to the automatic land cover mapping at the national scale from time series sets of multispectral satellite images may not only depend on the selected classifier. As land cover dynamics is a time continuous evolution without frequent short time variations, it is expected the correlation between time sequential multispectral images and vegetation indices to be high (Borak and Strahler 1999, Maxwell *et al.* 2002, Lu and Weng 2007). Indeed, the use of high frequency time series can be a source of noise for automatic classification and of high processing time without effective gains in classification accuracy (Maxwell *et al.* 2002, Knight *et al.* 2006). Moreover, the combination of multispectral information from indiscriminate dates may decrease the classification accuracy instead of improving it (Knight *et al.* 2006, Lu and Weng 2007). Thus, effective ways to reduce the dimensional feature space of these remotely-sensed data sets, while maintaining the key information for discriminating between cover types, is imperative and must be considered to easier the automatic land cover classification processes (Haertel and Landgrebe 1999, Landgrebe 2002, Guyon and Elissee 2003).

To tackle this problem, i.e. the dimensional feature space reduction, we focus on the analysis of the contribution of multitemporal and multispectral information from satellite images for land cover classes' discrimination. In the sequence, we pursue the exploitation of multitemporal images and investigate a mathematical function to methodically describe the intra-annual evolution of multispectral data and vegetation indices for disparate land cover classes. The goal is to optimize time information usage for the characterization of land cover distribution and land cover monitoring. Parametric models, like the one we propose, perform a *dimension reduction* to meet a trade-off between the searching space dimension and the number of classes to be differentiated, and optimize the classification performances (Bradley *et al.* 2007).

1.3.3. The line of research

Several authors have already focused on the problem of automatic land cover mapping from multitemporal and multispectral satellite images of medium spatial resolution (e.g. Borak and Strahler 1999, Maxwell *et al.* 2002, Clevers *et al.* 2007, Dash *et al*. 2007). However, the differences in the selected land cover nomenclatures, the used classification approaches, the attained land cover products and respective thematic accuracies highlight the potential for further work in this field.

In this research, we explore the use of multitemporal information from multispectral satellite images for land cover characterization. In detail, one investigates the usefulness of time series sets of medium spatial resolution satellite images for the regular land cover mapping and land cover monitoring at the national scale. To tackle this issue, we follow a rigorous process in three main steps: 1) assessment of the multitemporal information contribution for land cover classes discrimination at the national scale; 2) definition of an exhaustive land cover mapping protocol for the automatic production of a land cover map with a set of multitemporal satellite images and respective accuracy assessment; and 3) development of a model to fit time series sets of multispectral satellite images for optimizing time information usage for the characterization of land cover distribution and land cover monitoring.

This study focus on satellite images time series acquired by MODIS and MERIS sensors over the Earth's surface. The selected study area is the Portuguese mainland territory, located in the South-West area of Europe. The spatial resolution of the satellite images under analysis are well adapted to the landscape characteristics of the study area, and these are sufficiently diversified to test different methodological approaches for the characterization of land cover types with distinct biophysical properties, i.e. Atlantic and Mediterranean.

To assist the comprehension of the whole research work, next we briefly describe the main research points of each stand-alone manuscript that constitutes the structure of this dissertation. In the first study of this research (Chapter 2), we focus on the assessment of the benefits of multitemporal information from multispectral satellite images for land cover classes' discrimination in the framework of automatic classification. To tackle this issue and for evaluating the land cover classifications performed with single and multiple dates' multispectral information, we use a Support Vector Machine (SVM) learning approach and a cross validation technique.

Because the multispectral information contained in an intra-annual time series set of satellite images is correlated, in the second manuscript (Chapter 3) we evaluate the separated contribution of both the temporal and spectral information for land cover classes'

discrimination. In order to make this evaluation independent of some specific classifier, we propose the use of a metric based on the median Mahalanobis distance for the analysis of land cover classes' separability on different temporal and spectral dimensional features spaces.

In the sequence, and after the analysis of the contribution of the multitemporal and multispectral satellites images information for land cover classes discrimination, in Chapter 4 we focus on the selection of the best medium spatial resolution satellite images, i.e. MERIS or MODIS, for the land cover characterization in Continental Portugal. In Chapter 5, we compare the automatic land cover classification accuracy attained with several standard supervised classifiers with that attained with a machine learning approach, i.e. the Support Vector Machines (SVM). The goal is to select the best classification technique to automatically discriminate between the land cover classes of interest at the national scale, by using the spectral and temporal information of the elected medium spatial resolution satellite images. Moreover, in Chapter 6 we compare these standard classifiers with an Artificial Neural Network (ANN) classifier, i.e. the Self-Organizing Maps (SOM).

In Chapter 7, we propose a probabilistic sampling design that aims at improving the testing sample collection and the accuracy estimators' computation in the framework of a regular land cover mapping program. Afterwards, and following the specificities of the proposed sampling design, we collect a land cover sample database for the accuracy assessment of medium spatial resolution land cover maps within the Portuguese territory (Chapter 8). Within this study, we also propose the collection of ancillary information, namely a second reference label and nominal interpretation and location confidence ratings for each sample individual, to help in the posterior assessment of land cover maps thematic accuracy.

Following the outcomes of the previous stand alone manuscripts, we define in Chapter 9 a complete classification protocol and proceed with the automatic production of a land cover map with 16 classes for the study area of Continental Portugal. To evaluate the overall and specific thematic accuracies of this map, we use the testing sample database collected for the Portuguese territory.

To finish this investigation and to link our outcomes with future research, we propose in Chapter 10 a nonlinear harmonic model for fitting time series sets of multispectral reflectances and vegetation indices derived from satellite images for disparate land cover classes. The goal is to enhance the use of time series sets of multispectral reflectances and vegetation indices for the characterization of land cover distribution and time evolution in the framework of real or near-real time applications. We believe this is an innovative

approach for land cover characterization that simplifies the analysis of time series sets of satellite images by non-expert users. Moreover, this study preserves the continuity of this research, whereas allowing for posterior theoretical developments in the framework of remote sensing time series based applications.

1.4. References

Anderson, J.F., Hardy, E.E., Roach, J.T. and Witmer, R.E. (1976). *A land use and land cover classification system for use with remote sensor data*. U.S. Geological Survey Professional Paper 964, U.S. Geological Survey: Washington, DC, 28 pp.

Bartholomé, E. and Belward, A.S. (2005) GLC2000: a new approach to global land cover from Earth observation data. *International Journal of Remote Sensing*, 26, pp. 1959-1977.

Borak, J.S. and Strahler, A.H. (1999). Feature selection and land cover classification of a MODIS-like data set for a semiarid environment. *International Journal of Remote Sensing*, 20, 919-938.

Bradley, B.A., Jacob, R.W., Hermance, J.F. and Mustard, J.F. (2007). A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite NDVI data. *Remote Sensing of Environment*, 106, 137–145.

Caetano, M., Freire, S. and Carrão, H. (2004). Fire risk mapping by integration of dynamic and structural variables. In R. Goossens, (Ed), *Remote Sensing in Transition* (pp. 319-326). Roterdam: Millpress.

Caetano, M. and Araújo, A. (2006). Comparing land cover products CLC2000 and MOD12Q1 for Portugal. In A. Marçal (Ed.), *Global Developments in Environmental Earth Observation from Space* (pp. 469-477). Rotterdam: Millpress.

Carrão, H. (2002). *Caracterização da Estrutura da Paisagem: Modelação e Avaliação das Transformações na Representação da Ocupação do Solo*. Trabalho Fim de Curso da Licenciatura em Engenharia Biofísica, Universidade de Évora, 138 pp.

Cerdeira, C., Araújo, A., Carrão, H. and Caetano, M. (2006). Validação das Cartografias Globais de Ocupação do Solo, GLC2000 e MOD12Q1, para Portugal Continental. In *Actas do IX Encontro de Utilizadores de Informação Geográfica* (unpaginated CD-ROM), 15 – 17 November 2006, Oeiras, Portugal.

Cihlar, J. (2000). Land cover mapping of large areas from satellites: status and research priorities. *International Journal of Remote Sensing*, 21, 1093-1114.

Comber A.J. (2008). Land Cover or Land Use? *Journal of Land Use Science*, 3, 199–201.

Clevers, J. G. P. W., Schaepman, M. E., Mücher, C. A., de Wit, A. J. W., Zurita-Milla, R. and Bartholomeus, H. M. (2007) Using MERIS on Envisat for land cover mapping in the Netherlands. *International Journal of Remote Sensing*, 28, 637-652.

Cohen, W.B. and Goward, S.N. (2004). Landsat's Role in Ecological Applications of Remote Sensing. *Bioscience*, 54, 535-545.

Curran, P.J. and Steele, C.M. (2005). MERIS: the re-branding of an ocean sensor. *International Journal of Remote Sensing*, 26, 1781-1798.

Dash, J., Mathur, A., Foody, G.M., Curran, P.J., Chipman, J.W. and Lillesand, T.M., (2007). Land cover classification using multi-temporal MERIS vegetation indices. *International Journal of Remote Sensing*, 28, 1137-1159.

Di Gregorio, A. and Jansen, L.J.M. (2000). *Land Cover Classification System*. Rome. Italy: FAO, 179 pp.

EEA - European Environment Agency (2002). I&CLC2000 Technical Reference Document. Joint Research Centre, European Comission. Available online at: http://terrestrial.eionet.eu.int.

Foley *et al*. (2005). Global Consequences of land use. *Science*, 309, 570-574.

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

Foody, G. and Mathur, A. (2004). Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification. *Remote Sensing of Environment*, 93, 107−117.

Foody, G. M. and Mathur, A. (2006). The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. *Remote Sensing of Environment*, 103, 179-189.

Friedl, M.A., McIver D.K., Hodges J.C.F., Zhang X.Y., Muchoney D., Strahler A.H., Woodcock, C.E., Schneider A., Cooper, A., Baccini, A., Gao, F. and Schaaf, C. (2002). Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment*, 83, 287-302.

Giri, Z., Zhu, Z. and Reed, B. (2005). A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets. *Remote Sensing of Environment*, 94, 123-132.

Guyon, I. and Elissee, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3, 1157-1182.

Haertel, V. and Landgrebe, D. (1999). On the Classification of Classes with Nearly Equal Spectral Responses in Remote Sensing Hyperspectral Image Data. *IEEE Transactions on Geoscience and Remote Sensing*, 37, 2374-2386.

Hansen M.C., Defries R.S., Townshend J.R.G. and Sohlberg R. (2000). Global land cover classification at 1km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, 21, 1331-1364.

Hermance, J. F., Jacob, R. W., Bradley, B. A. and Mustard, J. F. (2007). Extracting Phenological Signals from Multi-Year AVHRR NDVI Time Series: Framework for Applying High-Order Annual Splines with Roughness Damping. *IEEE Transactions on Geoscience and Remote Sensing*, 45, 3264-3276.

Hodges, J. (2001). MODIS MOD12 Land Cover and Land Cover Dynamics Products User Guide. Available at: http://geography.bu.edu/landcover/userguidelc/ intro.html (last accessed, 3 June, 2005).

Homer, C., Huang, C., Yang, L., Wylie, B. and Coan, M. (2004). Development of a 2001 National Landcover Database for the United States. *Photogrammetric Engineering and Remote Sensing*, 70, 29-840.

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

Hughes, G. F. (1968). On the mean accuracy of statistical pattern recognizers. *IEEE Transactions on Information Theory*, 14, 55-63.

IGP – Instituto Geográfico Português (2009). Carta de Ocupação do Solo - COS' 90. Available at: http://www.igeo.pt/produtos/CEGIG/COS.htm (last accessed, 8 May, 2009).

Jackson, Q. and Landgrebe, D. A. (2001). An Adaptive Classifier Design for High-Dimensional Data Analysis with a Limited Training Data Set. *IEEE Transactions on Geoscience and Remote Sensing*, 39, 2664-2679.

Kalnay, E. and Cai, M. (2003). Impact of urbanization and land-use change on climate. *Nature*, 423, 528-531.

Kaufmann, R.K. and Seto, K.C. (2001). Change detection, accuracy, and bias in a sequential analysis of Landsat imagery: econometric techniques. *Agriculture, Ecosystems and Environment*, 85, 95–105.

Knight, J.F., Lunetta, R.L., Ediriwickrema, J. and Khorram, S. (2006). Regional Scale Land-Cover Characterization using MODIS-NDVI 250 m Multi-Temporal Imagery: A Phenology Based Approach. *GIScience and Remote Sensing*, 43, 1-23.

Lambin, E.F., Rounsevell, M. and Geist, H. (2000). Are current agricultural land use models able to predict changes in land use intensity?. *Agriculture, Ecosystems and Environment*, 82, 321-331.

Landgrebe, D. (2002). Hyperspectral Image Data Analysis as a High Dimensional Signal Processing Problem. *Special Issue of the IEEE Signal Processing Magazine* (Invited), 19, 17-28.

Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L. and Merchant, J.W. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, 21, 1303–1330.

Lu, D. and Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 28, 823-870.

Maxwell S. K., Hoffer, R. and Chapman, P. L. (2002). AVHRR composite period selection for land cover classification. *International Journal of Remote Sensing*, 23: 5043-5059.

Mathur, A. and Foody, G.M. (2006). Crop classification by a SVM with intelligently selected training data for an operational application. *International Journal of Remote Sensing*, 29, 2227–2240.

Meyer, W. B. and Turner II, B.L. (1994).Global Land-Use and Land-Cover Change: An Overview. In W. B. Meyer and B. L. Turner II (Eds.), *Changes in land use and land cover: a global perspective* (pp. 3– 10). Cambridge: Cambridge University Press.

Meyer, D., Leisch, F. and Hornik, K. (2003). The support vector machines under test. *Neurocomputing*, 55, 169-186.

Moran, E.F., Skole, D.L: and Turner II, B.L. (2004). The Development of the International Land Use and Land Cover Change (LUCC) Research Program and Its Links to NASA's Land Cover and Land Use Change (LCLUC) Initiative. In Gutman *et al.* (Eds.), *Land Change Science: Observing, Monitoring, and Understanding Trajectories of Change on the Earth's Surface* (pp. 1-16). New York: Kluwer Academic Publishers.

Moulin, S., Dondeau, A. and Delecolle, R. (1998). Combining agricultural crop models and satellite observations from field to regional scales. International Journal of Remote Sensing, 19, 1021–1036.

Mücher, C.A., Steinnocher, K.T., Kressler, F.P. and Heunks, C. (2000). Land cover characterization and change detection for environmental monitoring of pan-Europe. *International Journal of Remote Sensing*, 21, pp. 1159–1181.

Muchoney. D. M., Borak, J., Chi, H., Friedl, M., Hodges, J., Morrow, N. and Strahler, A. (2000). Application of the MODIS Global Supervised Classification Model to vegetation and land cover mapping of Central America. *International Journal of Remote Sensing*, 21, 1115-1138.

Müller, K., Mika, S., Rätsch, G., Tsuda, K. and Schölkopf, B. (2001). An introduction to kernel-based learning algorithms. *IEEE Transactions on Neural Networks*, 12, 181-201.

Olesen, J.E. and Bindi, M. (2002). Consequences of climate change for European agricultural productivity, land use and policy. *European Journal of Agronomy*, 16, 239-262.

Pal, M. and Mather, P.M. (2003). An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment*, 86, 554-565.

Pal, M. and Mather, P.M. (2005). Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, 26, 1007-1011.

Penner, J. E. (1994). Atmospheric chemistry and air quality. In W. B. Meyer and B. L. Turner II (Eds.), *Changes in land use and land cover: a global perspective* (pp. 175– 209). Cambridge: Cambridge University Press.

Rast, M., Bézy, J.L. and Bruzzi, S. (1999). The ESA Medium Resolution Imaging Spectrometer MERIS — a review of the instrument and its mission. *International Journal of Remote Sensing*, 20, pp. 1681–1702.

Rogan, J. and Chen, D. (2004). Remote sensing technology for mapping and monitoring land-cover and land-use change.*Progress in Planning*, 61, 301-325.

Skole, D.L. and Cochrane, M.A. (2004). Observations of LCLUC in regional case studies. In Gutman *et al.* (Eds.), *Land Change Science: Observing, Monitoring, and Understanding Trajectories of Change on the Earth's Surface* (pp. 53-56). New York: Kluwer Academic Publishers.

Steffen *et al*. (1998). The terrestrial carbon cycle: Implications for The Kyoto Protocol. *Science*, 280, 1393–1394.

Stehman, S. (1996). Estimating the kappa coefficient and its variance under stratified random sampling, *Photogrammetric Engineering and Remote Sensing*, 62, 401–407.

Stehman, S. and Czaplewski, R. (1998). Design and analysis for thematic map accuracy assessment: fundamental principles. *Remote Sensing of Environment*, 64, 331–344.

Stehman, S., Czaplewski, R., Nusser, S., Yang, L. and Zhu, Z. (2000). Combining Accuracy Assessment of Land-Cover Maps with Environmental Monitoring Programs. *Environmental Monitoring and Assessment*, 64, 115-126.

Stehman, S. (2001). Statistical rigour and practical utility in thematic map accuracy, *Photogrammetric Engineering and Remote Sensing*, 67, 727–734.

Treitz, P. and Rogan, J. (2004). Remote sensing for mapping and monitoring land-cover and land-use change—an introduction. *Progress in Planning*, 61, 269-279.

Van Niel, T. G., McVicar, T. R. and Datt, B. (2005). On the relationship between training sample size and data dimensionality of broadband multi-temporal classification. *Remote Sensing of Environment*, 98, 468−480.

Vapnik, V. (1998). *Statistical Learning Theory*. New York: John Wiley.

Veldkamp, A. and Lambin, E.F. (2001). Predicting land use change, *Agriculture, Ecosystems and Environment*, 85, 1–6.

Verstraete, M.M., Pinty, B. and Curran, P.J.(1999). MERIS potential for land applications. *International Journal of Remote Sensing*, 20, 1747–1756.

Vitousek, P. M. (1994). Beyond global warming: ecology and global change. *Ecology*, 75, 1861– 1876.

Wils, W. P. J. (1994). The Birds Directive 15 years later: a survey of the case law and a comparison with the Habitats Directive. *Journal of Environmental Law*, 6, 219-242.

Zhan, X., Sohlberg, R. A., Townshend, J. R. G., Dimiceli, C., Carrol, M. L., Eastman, J. C., Hansen, M. C. and Defries, R.S (2002). Detection of land cover changes using MODIS 250 m data. *Remote Sensing of Environment*, 83, 336 – 350.

X. Y. Zhang *et al*. (2003). Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, 84, 471–475.

2 USE OF INTRA-ANNUAL SATELLITE IMAGERY TIME-SERIES FOR LAND COVER CHARACTERIZATION PURPOSE

Carrão, H., Gonçalves, P. and Caetano, M. (2007). Use of intra-annual satellite imagery time-series for land cover characterization purpose. *EARSeL eProceedings*, 6, 1-11.

Abstract

Automatic image classification often fails at separating a large number of land cover classes that punctually may present similar spectral reflectances. To improve classification accuracy in such situations, multi-temporal satellite data has proven to be valuable auxiliary information. In this paper we present a study exploring the usefulness of intra-annual satellite images time-series for automatic land cover classification. The reported work aims at producing a land cover classification of Continental Portugal from multi-spectral and multi-temporal MODIS satellite images acquired at a 500 meters spatial resolution for the year 2000. We started our study by performing a single date classification to define the month with the best score as a benchmark to compare with classification accuracies obtained with sets of images from various dates. Then, we considered various combinations of twelve intra-annual image observations (one per month) to quantify the gain when integrating temporal information in the classification process. Curiously, the results we obtained show that multi-temporal information does not significantly improve overall classification accuracy, but in particular it permits to better separate similar land cover classes even if those remain wrongly identified. Surprisingly also, we show that only few (typically 2) dates are sufficient to reach optimal performance of our multi-temporal classifier. In our study we used a Support Vector Machine learning approach.

Keywords: MODIS, intra-annual time-series, land cover, Support Vector Machine.

2.1. Introduction

Remotely sensed images of the Earth's surface have been widely used in the past decades for deriving land cover information by means of automatic classification. Commonly, land cover mapping involves single date image analyses. However, since maximum discrimination between different land cover classes occurs at different stages in the growth cycle of vegetation types, this approach has the drawback that not all differences are incorporated in the procedure (1). Thus, mapping of land cover often requires processing satellite images

collected at different time periods and at many spectral wavelengths (2, 3). Still, it is not clear yet which temporal observations are required to completely characterize land cover types, nor the specific effects of annual climatic anomalies on the selection of critical dates for land cover characterization (4). Nevertheless, multi-temporal satellite image datasets provide valuable information on the phenological characteristics of vegetation, thereby increasing the accuracy of cover type classifications compared to single date classifications (5, 6).

Remote sensors capable of providing surface information over large regions have usually a high temporal resolution but coarse spatial and spectral resolutions. Until recently, the Advanced Very High Resolution Radiometer (AVHRR), with 1.1 km spatial resolution and daily temporal resolution (7, 8), was one of the few sensors on orbit with such properties. Thus, several available regional and global land cover classification studies and operational programs were based on AVHRR data, see e.g. (9) and (10), to cite but a few. In these situations, the multi-temporal spectral information provide a valuable substitute for the enhanced spectral and spatial characteristics of high resolution remote sensors (e.g., Landsat and SPOT), which have been mainly exploited for automatic land cover classification at national and local scales.

Nowadays, other Earth Observation (EO) sensors with high temporal resolutions, such as the MEdium Resolution Imaging Spectrometer (MERIS) and the Moderate Resolution Imaging Spectroradiometer (MODIS) are also available, featuring better spatial resolutions (up to 300 and 250 m, respectively) and superior standards of calibration, georeferencing and atmospheric correction, as well as detailed per pixel data quality information. As such, the EO community has started to explore images acquired by these two sensors for the land cover assessment at regional and global scales (e.g. 11, 12, 13, 14, 15, 16, 17, and 18). In particular, spectral, temporal, and spatial information acquired by these sensors are all included as part of the feature space exploited for land cover classification of wide regions (19). Imagery time-series analysis remains an essential approach for land cover cartography production at medium spatial scales, although high spectral resolution imagery allows for retrieving more detailed land cover features with single date information than before (3).

In this study we compare the accuracy of land cover classifications performed in Portugal with single date and intra-annual MODIS images time-series acquired at a nominal resolution of 500m. Specifically, we evaluate if classification accuracy is significantly improved with multi-temporal spectral information in opposition to single date spectral data. Land cover classification results were computed with Support Vector Machine (SVM), a

recently proposed supervised classification system that is insensitive to space dimensionality (20).

2.2. Study area and data

The study area is the entire Portuguese mainland territory. Portugal is in a transition zone featuring diverse landscapes representing both Mediterranean and Atlantic climate environments. This landscape heterogeneity allows for the extrapolation, to some extend, of the developed methodologies to other regions of the world.

Our study relies on the MOD09A1 product, an 8 days composite of surface reflectance images, freely available from MODIS Data Product web site (http://modis.gsfc.nasa.gov). We considered a full year observation period, from February 2000 to January 2001 (12 cloud free images), of surface reflectances measured within seven disjoint spectral bands (VIS+SWIR+MIR) and imaged at a nominal spatial resolution of 500 meters. Moreover, two vegetation indices (i.e. *Normalized Difference Vegetation Index* – NDVI (Equation (2.1); *Enhanced Vegetation Index* – EVI (Equation (2.2) were also calculated for each date and used as additional band information.

$$NDVI = (\rho_{nir} - \rho_{red})/(\rho_{nir} + \rho_{red}) \quad (2.1)$$

$$EVI = 2.5\frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + 6\times\rho_{red} - 7.5\rho_{blue} + 1} \quad (2.2)$$

where ρ_{nir}, ρ_{red} and ρ_{blue} represent the surface reflectance of near-infrared (B2), red (B1) and blue (B3) MOD09A1 bands, respectively.

The land cover classes used in this study are the following: Water Bodies (WB), Natural grassland (Ng), Broadleaved Closed Trees (BCT), Barren (B), Shrubland (S), Needleleaved Closed Trees (NCT), Irrigated Herbaceous Crops (IHC), Rain Fed Herbaceous Crops (RHC) and Continuous Artificial Areas (CAA).

2.3. Methodology

2.3.1. Sampling

In order to test the several classification approaches using the defined nomenclature, we selected a collection of representative sample units for the nine classes of the nomenclature for the year 2000. Each sample unit represents a specific land cover class covering a surface area of 500-by-500 m (same as the nominal resolution of used satellite images) and the final set was acquired following a stratified random sampling design using the CORINE Land Cover 2000 (CLC2000) cartography (21) as strata. To recheck every sample unit collected

for each land cover class and take into account CLC2000 map generalization procedure we used ancillary data (e.g. Landsat ETM+ images acquired during the year of 2000 and orthorectified colour infrared aerial photography of 1995). At the end of this process 354 sample units were uniformly collected all over the mainland territory and distributed among the classes as follows: WB-40, Ng-16, BCT-40, B-27, S-18, NCT-55, IHC-52, RHC-59, and CAA-47.

2.3.2. Classification

Land cover classification results were computed with Support Vector Machine (SVM). SVM are a new generation of supervised learning systems based on recent advances in statistical learning theory (22). Pioneered by the work on learning strategy by Vapnik and collaborators (23, 24), they have rapidly and successfully been applied to numerous real-world classification problems. In short, SVM is a kernel-based classifier that uses a non linear mapping to transpose the data initially lying in a non linearly separable space, onto a (possibly infinite dimension) feature space. Due to the high complexity of this new representation, it is likely that the different classes become linearly separable. In our specific task, we used Gaussian kernels and a regularization strategy that is often referred to as the ν-parameterization in the SVM literature (25). The generalization to multiple classes' problem was straightforward using a one-versus-the-rest strategy. See (11) and (13) for a complete and detailed description of the SVM learning system used in this study.

Twelve classifications with SVM, each one using a single-date image from a different month, were initially performed in order to compare the results of posterior multi-temporal images classifications against the best classification accuracy based on single date spectral information. In a second step, land cover classifications with SVM were performed with several combinatorial subsets of the twelve monthly images. We compute all possible classifications for subsets of two and three dates (66 and 220 combinations, respectively). A classification based on the set comprising all the twelve image dates was also performed. Other combinatorial subsets were discarded because we achieved the essential results with these arrangements.

There is little guidance in the literature on the criteria to be used in selecting the optimal kernel-specific parameters for SVM computation, so a number of trials were carried out with each dataset using different kernel-specific parameters, and taking into account classification accuracy as a measurement of their quality (20).

All classification results have been obtained using cross validation – a standard procedure used in longitudinal data analysis, e.g. land cover classification (26, 27, 28), that permits

evaluating the first and second order statistics of the classifier. Cross validation permits evaluating the performance of a classifier from a single data set used for train and test purposes simultaneously. The data is split into *K* subsets of samples, and *K* classifications are performed using each subset as a hold-out set. The classifier accuracy is scored on the hold-out set of each classification. The average accuracy is taken as the estimated accuracy over the entire domain that is to be classified. We fixed to five the number of cross validation folds (26).

2.3.3. Analysis

In order to perform a systematic investigation of the relative gain from incorporating a temporal dimension into land cover classification process, standard accuracy measures were derived from the classification results obtained with each combinatorial subset of the twelve monthly images. The measures used were: overall accuracy, producer's accuracy and user's accuracy (29).

Moreover, since the comparison of classifications was fundamental for the study, a statistically rigorous approach for asserting significant differences between our experimental setups was adopted (26). The goal was to evaluate $H_0 : p_a = p_b$ versus $H_1 : p_a < p_b$, in which p_a and p_b indicate, respectively, the proportion of correctly classified pixels from *a* and *b* images datasets, and moreover card (*b*) > card (*a*). The hypothesis testing for the difference between two variables with binomial sampling distribution is based upon the standardized normal test statistic (30, 31, 32):

$$Z = \frac{\hat{p}_a - \hat{p}_b}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{m}+\frac{1}{n}\right)}} \xrightarrow{a} N(0,1) \qquad (2.3)$$

in which $\hat{p}_a - \hat{p}_b$ is the sample estimate of $p_a - p_b$, $\hat{p} = \frac{m\hat{p}_a + n\hat{p}_b}{m+n}$, and *m* and *n* the number of sample observations used to derive $\hat{p}_a$ and $\hat{p}_b$, respectively. With this test, a significant increase in classification accuracy occurs at 95% level of confidence if $Z < (-1.645)$.

2.4. Results and discussion

In Figure 2.1 we present the overall 9-classes classification accuracy obtained with SVM applied to each of the monthly dates separately. The goal was to define the month yielding the best score and to use this as a benchmark to position the classification accuracies obtained with multi-temporal image sets.

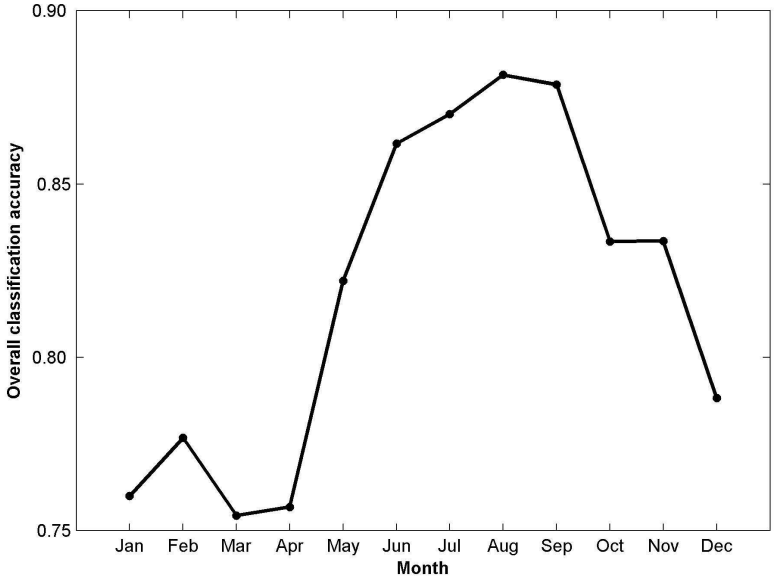

Figure 2.1. Overall classification accuracy as a function of time.

The best classification rates are attained in summer times (between June and September), with August the month with the maximum overall classification accuracy. These results are similar to the ones obtained by (13), confirming that this is the best period to discriminate between core cover types in Portugal.

A closer view at the user's and producer's accuracies per class (Table 2.1), based on the classification results derived with the reflectances and vegetation indices measured in August image, shows that "Natural Grassland" (Ng) is definitely the least distinguishable land cover class, frequently confounded with "Shrubland" (S), "Barren" (B) and "Rain Fed Herbaceous Crops" (RHC). Similarly, "Shrubland" (S) is often confounded with several land cover classes, but mainly with "Natural Grassland" (Ng).

Looking at Figure 2.2 we perceive that the averaged EVI time series corresponding to "Barren" (B) and "Natural grassland" (Ng) overlap in the entire summer period. Furthermore, the "Natural grassland" (Ng) profile is up to a scale factor, very similar to the "Rain Fed Herbaceous Crops" (RHC) response. Similar results were observed with NDVI and spectral bands. For these reasons, it is not surprising that classification results computed with a single summer image reveal a mixture between these cover types. However, and given the different phenology of these classes, it is reasonably expectable that adding more temporal information in the feature space, selected at phenologically critical times (e.g. February or November), should improve their particular classification performances (6). The

complexity is to define which summer and winter dates should be combined to increase land cover classes discrimination.

Table 2.1. User's and Producer's accuracies per class. This repartition corresponds to the SVM classification based on reflectances and vegetation indices obtained at a single date (August, 2000).

	WB	**Ng**	**BCT**	**B**	**S**	**NCT**	**IHC**	**RHC**	**CAA**	**User's accuracy**
WB	40	0	0	0	1	0	0	0	0	0.98
Ng	0	5	0	1	4	0	0	2	0	0.42
BCT	0	0	40	0	2	2	0	0	0	0.91
B	0	4	0	18	2	0	1	0	1	0.69
S	0	3	0	2	6	0	1	0	0	0.50
NCT	0	0	0	0	2	53	0	0	0	0.96
IHC	0	0	0	0	1	0	50	0	0	0.98
RHC	0	4	0	1	0	0	0	56	0	0.92
CAA	0	0	0	5	0	0	0	1	46	0.88
Producer's accuracy	1.00	0.31	1.00	0.67	0.33	0.96	0.96	0.95	0.98	-

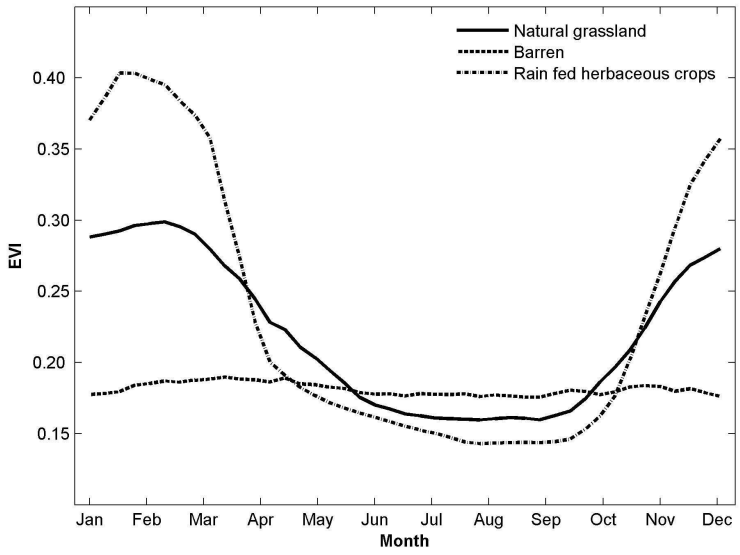

Figure 2.2. Mean EVI values as function of time for "Natural grassland" (Ng), "Barren" (B), and "Rain Fed Herbaceous Crops" (RHC). See (33) for further details on mean profiles calculation.

In Table 2.2 we present the results of the classifications performed with combinatorial subsets of two dates. In all cases, a SVM is trained and optimized on reflectance from all spectral bands and vegetation indices. In Figure 2.3 we present, for each month m_i

(i=1,…,12), a box plot with the overall classification accuracy statistics (median, quartile, extremes and outliers) obtained for all pair-wise combinations of m_i with each of the remaining eleven months (Table 2.2).

Table 2.2. Overall land cover classification accuracies obtained with combinatorial subsets of two dates.

	Jan	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
Jan	-	0.84	0.84	0.85	0.88	0.88	0.89	0.89	0.88	0.86	0.85	0.84
Feb	0.84	-	0.81	0.85	0.86	0.87	0.87	0.88	0.87	0.86	0.86	0.85
Mar	0.84	0.81	-	0.81	0.86	0.87	0.88	0.88	0.88	0.87	0.86	0.85
Apr	0.85	0.85	0.81	-	0.87	0.87	0.87	0.88	0.87	0.87	0.87	0.86
May	0.88	0.86	0.86	0.87	-	0.86	0.87	0.87	0.88	0.87	0.87	0.87
Jun	0.88	0.87	0.87	0.87	0.86	-	0.86	0.88	0.87	0.89	0.89	0.88
Jul	0.89	0.87	0.88	0.87	0.87	0.86	-	0.87	0.88	0.89	0.90	0.89
Aug	0.89	0.88	0.88	0.88	0.87	0.88	0.87	-	0.87	0.88	0.90	0.89
Sep	0.88	0.87	0.88	0.87	0.88	0.87	0.88	0.87	-	0.87	0.88	0.88
Oct	0.86	0.86	0.87	0.87	0.87	0.89	0.89	0.88	0.87	-	0.86	0.87
Nov	0.85	0.86	0.86	0.87	0.87	0.89	0.90	0.90	0.88	0.86	-	0.85
Dec	0.84	0.85	0.85	0.86	0.87	0.88	0.89	0.89	0.88	0.87	0.85	-

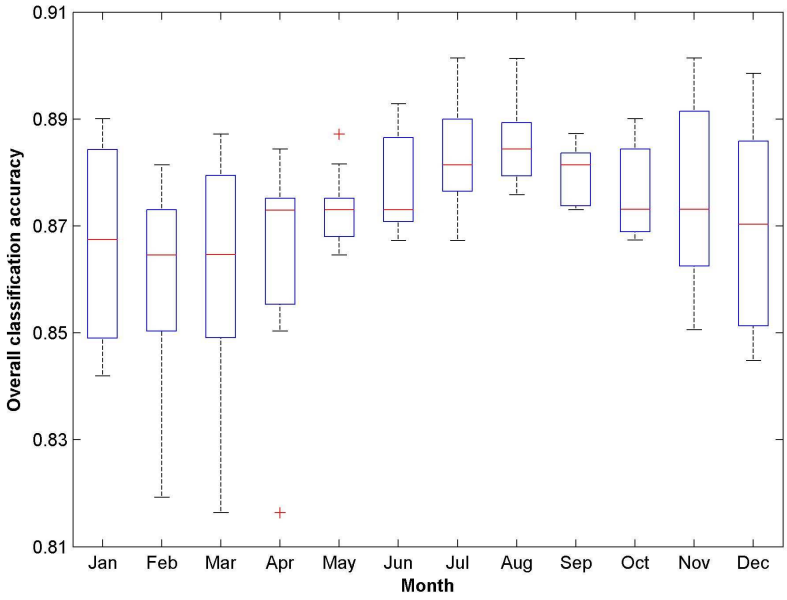

Figure 2.3. Overall classification accuracy per combinatorial subsets of two dates. Each box plot represents the lower quartile, median, and upper quartile of classification accuracy obtained for all pair-wise combinations of each month with each of the remaining eleven months images. Classification values that are considered as outliers are represented by "+".

The most important remarks when comparing these results with those from Figure 2.1 is that the best median pair-wise classifications are achieved considering at least one image from summer time, and overall classification accuracies with summer images are those that have the smallest intervals of variation. These outcomes confirm, once again, that spectral information from summer is necessary for superior land cover classes' discrimination in Portugal. Specifically, the maximum overall classification accuracy was obtained by combining August and November images (the same value was achieved with the subset containing July and November dates). Indeed, the median values of the classifications achieved by combining August with all other months proved to be greater than the median accuracy achieved with pair-wise combinatorial sets for all other dates.

This result suggests that is most likely to obtain a maximum overall classification accuracy value if spectral information from August is in the dataset used for land cover classification. Moreover, looking closer at the results from Figure 2.2, we also perceive that November and December are some of the winter dates that maximize the distance between the mean profiles of cover types that mix up in summer (e.g., Ng, and B). For that reason, it is expected that the classification accuracies of these classes could be increased by using these dates together with the summer dates as input features for classification. Classification results with two dates proved that the previous analysis of EVI mean profiles of Figure 2.2 is sensitive, and that the overall and specific classification accuracies can be increased by using spectral features from different seasons (summer and winter) in the classification process (Table 2.3).

A similar classification approach with combinatorial subsets of three dates was also performed. Figure 2.4 compares the best overall classification accuracies obtained when training a SVM (i.e. optimizing its parameters) on observations from month m_i, from the pair (m_i, m_j), and from the triplet (m_i, m_j, m_k). For each m_i-combination, only the best classification accuracy is displayed.

Figure 2.4 shows that increasing the temporal spectral features for classification maximizes the monthly overall classification accuracy achievable with single date information. However, this is mostly evident for months out of the summer period, which are the ones that clearly benefit from those combinations, whereas it is not so important for the classification results achieved with a single date in summer. This is an important result showing that the best classification accuracy attained with single date information is quite similar to those attained with two and three dates.

Classification accuracies for the nine land cover classes using single-date, two-dates, three-dates and twelve-dates spectral information and vegetation indices as input features for the

classification with SVM are presented in Table 2.3. Notice that for two and three dates, only the combinatorial set with the best overall classification accuracy in each case is presented.

Table 2.3. Overall classification accuracies from the best single date classification, from the best two and three dates combinatorial sets and from the twelve dates association. The table shows the individual classification accuracy for each class (producer's accuracy) and overall accuracy.

	Number of input dates for classification			
Land cover class	**1 (August)**	**2 (August + November)**	**3 (August + November + July)**	**12 Dates**
Water Bodies (WB)	1.00	1.00	1.00	1.00
Natural grassland (Ng)	0.31	0.31	0.31	0.38
Broadleaved Closed Trees (BCT)	1.00	0.98	0.95	0.95
Barren (B)	0.67	0.70	0.74	0.67
Shrubland (S)	0.33	0.44	0.39	0.33
Needleleaved Closed Trees (NCT)	0.96	0.96	0.96	0.96
Irrigated Herbaceous Crops (IHC)	0.96	0.98	1.00	0.98
Rain Fed Herbaceous Crops (RHC)	0.95	0.98	1.00	0.98
Continuous Artificial Areas (CAA)	0.98	0.98	0.98	1.00
OVERALL ACCURACY	0.89	0.90	0.90	0.90

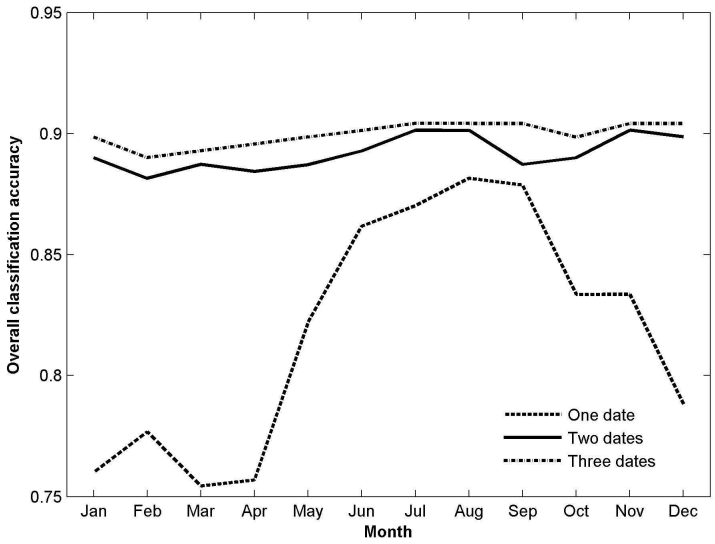

Figure 2.4. Overall classification accuracy obtained with observations from a single date and from combinatorial sets of two and three dates. For each month, only the best overall classification accuracies obtained by combining it with one and two other dates are displayed.

Looking at the results of Table 2.3 we perceive that overall and particular classification accuracies obtained with one or more dates are quite similar. The classification results showed that embedding information available for all dates simultaneously had to face the problem of insufficient training data, which modestly degraded the classification scores previously obtained for some classes with only two and three dates. Also, the sparseness of available samples adds to the classification difficulty. However, we notice that the overall accuracy, as well as the producer's accuracies of "Shrubland" (S), "Barren" (B) and agricultural (IHC, RHC) classes, were improved by using spectral data and vegetation indices from August and November images as input features for classification. Indeed, this is the most important result proving that multi-temporal data may be redundant regarding the overall classification task, but can show useful at separating two specific classes (which does not necessary mean that individually those classes will be correctly identified in a multi-class classification context).

Table 2.4 provides the *Z* values calculated in order to check how significant the improvements in overall classification accuracy are when we take into account more than one single-date spectral reflectances and vegetation indices for classification. *Z* values were computed considering the proportion of correctly classified sample units per combinatorial subset presented in Table 2.3. If the calculated *Z* value is smaller than -1.645, then the difference between the overall classification accuracies obtained with two different imagery time sets can be interpreted as a significant increase in classification accuracy at a 95% confidence level.

Table 2.4. *Z* values to appraise the significance of the difference between classification accuracies from the best single date classification, from the best two and three dates combinatorial sets and from the twelve dates association.

Number of input dates for classification	**1**	**2**	**3**	**12**
1	-	-0.61	-0.74	-0.36
2	-0.61	-	-0.13	0.25
3	-0.74	-0.13	-	0.38
12	-0.36	0.25	0.38	-

The results introduced in Table 2.4 show that there is no significant improvement between the performances of SVM using spectral data from one, two, three or even twelve dates as input features for classification (all calculated *Z* values are superior to -1.645). It can be

concluded that for this land cover sample set, all combinations of time features are equivalent for classification with SVM. This result was not surprising, since (3) have previously shown that the temporal domain was not relevant for land cover classification with MODIS data in Arizona. Still, it is important to note that overall classification accuracy results always benefit from the addition of multi-temporal information, as observed in Table 2.3.

2.5. Conclusions

The results we obtained show that combining intra-annual monthly spectral information for land cover classes' discrimination in Portugal does not significantly improve the best overall classification attained with single date information. Spectral data from summer dates proved to be indispensable for most land cover classes' discrimination, but pair-wise classes discrimination benefit from the inclusion of spectral data of at least one winter date. Thus, we should remark that the use of multi-temporal imagery data for the discrimination between particular land cover classes proved to be important, although overall accuracy did not significantly benefit from it. Indeed, since spectral profiles of vegetation classes are in general smooth slowly time varying functions, it is expectable that few temporal spectral features are sufficient to assemble most of the discriminant information.

From a methodological viewpoint, SVM proved to be insensitive to the dimensionality of input data, given that similar overall land cover classification accuracy results were obtained with one and twelve monthly imagery spectral data.

2.6. Acknowledgements

This study was carried out in the framework of the project "LANDEO - User driven land cover characterisation for multi-scale environmental monitoring using multi-sensor earth observation data (PDCTE/MGS/49969/2003)" funded by "Programa Dinamizador das Ciências e Tecnologias para o Espaço" from "Fundação para a Ciência e Tecnologia", and from the "Announcement of Opportunity for the Utilisation of ERS and ENVISAT Data" from European Space Agency (ESA). Research by Hugo Carrão was founded by the "Fundação para a Ciência e Tecnologia" (SFRH/BD/18447/2004). This work was partially performed while Hugo Carrão was visiting INRIA Rhône-Alpes granted with a six month INRIA scholarship.

2.7. References

1 Vieira C A O, P M Mather & P Aplin, 2001. Multitemporal classification of agricultural crops using the spectral-temporal response surface. In: Proceedings of

the First International Workshop on the Analysis of Multi-Temporal Remote Sensing Images, edited by L Bruzzone & P Smits (University of Trento, Italy), 290-297.

2 Maxwell S K, R Hoffer & P L Chapman, 2002. AVHRR channel selection for land cover classification. International Journal of Remote Sensing, 23: 5061-5073.

3 Borak J S & A H Strahler, 1999. Feature selection and land cover classification of a MODIS-like data set for a semiarid environment. International Journal of Remote Sensing, 20: 919-938.

4 Loveland T R, J M Merchant, D O Ohlen & J F Brown, 1991. Development of a land- cover characteristics database for the conterminous U.S.. Photogrammetric Engineering and Remote Sensing, 57: 1453-1463.

5 Maxwell S K, R Hoffer & P L Chapman, 2002a. AVHRR composite period selection for land cover classification. International Journal of Remote Sensing, 23: 5043-5059.

6 Knight J F, R L Lunetta, J Ediriwickrema & S Khorram, 2006. Regional Scale Land-Cover Characterization using MODIS-NDVI 250 m Multi-Temporal Imagery: A Phenology Based Approach. GIScience and Remote Sensing, 43: 1-23.

7 Van Dijk A, S L Callis, C M Sakamoto & W L Decker, 1987. Smoothing vegetation index profiles: an alternative method for reducing radiometric disturbances in NOAA/AVHRR data. Phtogrammetric Engineering and Remote Sensing, 53: 1059-1067.

8 Duchemin B, D Guyon & J P Lagouarde, 1999. Potential and limits of NOAA-AVHRR temporal composite data for phenology and water stress monitoring of temperate forest ecosystems. International Journal of Remote Sensing, 20: 895-917.

9 Brown J, T Loveland, D Ohlen & A Zhu, 1999. The Global Land Cover Characteristis Database: The Users' Perspective. Photogrammetric Engineering and Remote Sensing, 65: 1069-1064.

10 Laporte N, S J Goetz, C O Justice & M Heinicke, 1998. A new land cover map of central Africa derived from multi-resolution, multi-temporal AVHRR data. International Journal of Remote Sensing, 19: 3537-3550.

11 Carrão, H, P Gonçalves & M Caetano, 2006. MERIS Based Land Cover Characterization: A Comparative Study. In: Proceedings of the ASPRS 2006 Annual Conference - Prospecting for Geospatial Information Integration, Reno, Nevada.

12 Oliveira P, P Gonçalves & M Caetano, 2005. Land cover time profiles from linear mixture models applied to MODIS images. In: Proceedings of the 31st International

Symposium on Remote Sensing of Environment, Saint Petersburg, Russian Federation.

13 Gonçalves P, H Carrão, A Pinheiro & M Caetano, 2006. Land cover classification with Support Vector Machine applied to MODIS imagery. In: Global Developments in Environmental Earth Observation from Space, edited by A Marçal (Rotterdam: Millpress, 2006), 517-525.

14 Clevers J G P W, R Zurita Milla, M Schaepman & H Bartholomeus, 2004. Using MERIS on Envisat for Land Cover Mapping. In: Proceedings of the 2004 Envisat & ERS Symposium, Salzburg, Austria (ESA SP-572, April 2005).

15 Gessner U, K P Gunther & S W Maier, 2004. Landcover/land use map of Germany based on MERIS full resolution data. In: Proceedings of the 2004 Envisat & ERS Symposium, Salzburg, Austria (ESA SP-572, April 2005).

16 Skinner L & A Luckman, 2003. Deriving Land Cover information over Siberia using MERIS and MODIS data. In: Proceedings of the MERIS User Workshop, Frascati, Italy (ESA SP-549, May 2004).

17 Duong N D, 2004. Land cover mapping of Vietnam using MODIS 500m 32-day global composites. In: International Symposium on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences, Hanoi, Vietnam.

18 Stibig H & T Bucha, 2005. Feasibility study on the use of medium resolution satellite data for the detection of forest cover change caused by clear cutting of coniferous forests in the northwest of Eurasia, Institute for Environment and Sustainability - Joint Research Centre ISPRA, Italy: European Commission.

19 Friedl M A, D K McIver, X Y Zhang, J C F Hodges, A Schnieder, A Bacinni, A H Strahler, A Cooper, F Gao, C Schaaf & W Liu, 2001. Global land cover classification results from MODIS. In: Geoscience and Remote Sensing Symposium 2001. IGARSS '01. IEEE 2001 International, Sidney, Australia, 733 – 735.

20 Pal M & P M Mather, 2005. Support vector machines for classification in remote sensing. International Journal of Remote Sensing, 26: 1007-1011.

21 Painho M & M Caetano, 2005. Cartografia de Ocupação do Solo, Portugal Continental, 1985-2000. Instituto do Ambiente, Lisboa, Portugal.

22 Cristianini N & J Shawe-Taylor, 2000. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods (Cambridge University Press) 204 pp.

23 Boser B E, I M Guyon & V N Vapnik, 1992. A training algorithm for optimal margin classifiers. In: Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory, edited by D Haussler, 144-152.

24 Vapnik V, 1998. Statistical Learning Theory (Wiley), 768 pp.

25 Scholkopf B, A J Smola, R C Williamson & P L Bartlett, 2000. New support vector algorithms. Neural Computation, 12:1207–1245.

26 Foody G M, A Mathur, C Sanchez-Hernandez & D S Boyd, 2006. Training set size requirements for the classification of a specific class. Remote Sensing of Environment, 104: 1-14.

27 Huang C, L S Davis & J R G Townshend, 2002. An assessment of support vector machines for land cover classification. International Journal of Remote Sensing, 23: 725-749.

28 Groves P & P Bajcsy, 2003. Methodology for Hyperspectral Band and Classification Model Selection. In: IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data, 27 October 2003, Washington DC.

29 Jensen John R, 1996. Introductory Digital Image Processing, A Remote Sensing Perspective (Prentice Hall, 2nd edition) 316 pp.

30 Pestana D D & S F Velosa, 2006. Introdução à Probabilidade e à Estatística (Fundação Calouste Gulbenkian, 2.ª edição) 1164 pp.

31 Murteira B, C S Ribeiro, J A Silva & C Pimenta, 2002. Introdução à Estatística (McGraw-Hill) 647 pp.

32 Johnson R A & D W Wichern, 1998. Applied Multivariate Statistical Analysis (Prentice Hall, 4th edition) 799 pp.

33 Pinheiro A, P Gonçalves, H Carrão & M Caetano, 2006. A Toolbox for multi-temporal analysis of satellite imagery. In: New Developments and Challenges in Remote Sensing, edited by Z Bochenek (Rotterdam: Millpress, in press).

3 CONTRIBUTION OF MULTISPECTRAL AND MULTITEMPORAL INFORMATION TO LAND COVER CLASSIFICATION FROM MODIS IMAGES

Carrão, H., Gonçalves, P. and Caetano, M. (2008). Contribution of multispectral and multitemporal information from MODIS images to land cover classification. *Remote Sensing of Environment*, 112, 986-997.

Abstract

The goal of this study is to evaluate the relative usefulness of high spectral and temporal resolutions of MODIS imagery data for land cover classification. In particular, we highlight the individual and combinatorial influence of spectral and temporal components of MODIS reflectance data in land cover classification. Our study relies on an annual time series of twelve MODIS 8-days composite images (MOD09A1) monthly acquired during the year 2000, at a 500 m nominal resolution. As our aim is not to propose an operational classifier directed at thematic mapping based on the most efficient combination of reflectance inputs – which will probably change across geographical regions and with different land cover nomenclatures – we intentionally restrict our experimental framework to continental Portugal. Because our observation data stream contains highly correlated components, we need to rank the temporal and the spectral features according not only to their individual ability at separating the land cover classes, but also to their differential contribution to the existing information. To proceed, we resort to the median Mahalanobis distance as a statistical separability criterion. Once achieved this arrangement, we strive to evaluate, in a classification perspective, the gain obtained when the dimensionality of the input feature space grows. We then successively embedded the prior ranked measures into the multitemporal and multispectral training data set of a Support Vector Machines (SVM) classifier. In this way, we show that, only the inclusion of the approximately first three dates substantially increases the classification accuracy. Moreover, this multitemporal factor has a significant effect when coupled with combinations of few spectral bands, but it turns negligible as soon as the full spectral information is exploited. Regarding the multispectral factor, its beneficence on classification accuracy remains more constant, regardless of the number of dates being used.

Keywords: Statistical separability analysis; MODIS intra-annual composite data; land cover classification; Support Vector Machines.

3.1. Introduction

Multitemporal satellite images composites are now of standard use in land cover classification of large areas at regional and global scales (Cihlar 2000). Considering in the one hand, that single date images may fail at discerning different land cover types that temporarily share similar spectral reflectance characteristics (Vieira *et al.* 2001, Liu *et al.* 2003); and recalling on the other hand, that most of the worldwide landmass is covered by vegetation with distinct phenologies, it is not senseless to believe that non-stationarities of image time series are valuable sources of information to get more accurate land cover maps (Knight *et al.* 2006). Data sets from the Advanced Very High Resolution Radiometer (AVHRR) have been used to map land cover types through monthly composites describing seasonal variations in photosynthetic activity of vegetation (DeFries and Belward 2000). However, mainly two of the five broad spectral bands sensed by AVHRR instrument are valuable for land observation, thus being insufficient to distinguish subtle differences in vegetation types with similar annual phenologies (Borak and Strahler 1999). Even so, land-sensing instruments that collect data at higher spectral resolutions, such as Landsat Thematic Mapper (TM), were generally not used to derive comparable regional products due to incomplete spatial coverage, infrequent temporal coverage with inevitable cloud contamination and the associated large data volumes not practicable in an automatic land cover classification context at such geographical scale (DeFries and Belward 2000).

Recently launched Earth Observation (EO) sensors exhibit enhanced spectral and radiometric resolutions, wide geographical coverage and improved atmospheric corrections, while preserving a temporal resolution comparable to that of AVHRR. This is precisely the case for the Moderate Resolution Imaging Spectroradiometer (MODIS) which is at the core of our study. Regarding spectral coverage, MODIS spans a broad spectrum with a large number of narrow bandwidth channels that constitute an advantage for a large scope of tasks and circumstances (Landgrebe 2005). In this direction, recent literature has shown that narrow bands can outperform broad bands in quantifying biophysical vegetation characteristics (Thenkabail *et al.* 2000). In addition, MODIS has on-board automatic positional calibrations that reduce the inherent AVHRR susceptibility to problems of misregistration in multitemporal analyses. All these new potentialities should help addressing the challenges of automatic land cover classification of wide geographical areas (Carrão *et al.* 2007), recalling that MODIS data already proved successful in a host of

regional applications requiring fine scale cover information: mapping of irrigated areas in the Ganges and Indus river basins (Thenkabail *et al.* 2005), production of a landform map of North Africa (Ballantine *et al.* 2005), regional land cover mapping to monitor biological conservation in the Great Yellowstone Ecosystem (Wessels *et al.* 2004), land cover classification over the Yellow River basin, China (Matsuoka *et al.* 2004), to cite but a few.

However, increasing the number of spectral channels and lengthening the data time series does not necessarily mean that the quantity of effective information for land cover classification is increasing proportionally. The information gain depends on the mutual independence of the co-occurring measures, and often all the necessary structure to discriminate between land cover classes lies in a low dimensional feature space (Landgrebe 2005). On the other hand, data sets with oversized dimension directly penalize the performances of most supervised classifiers. As the dimensionality increases with the number of spectral channels, the number of training samples for training a specific classier should be increased exponentially as well. This effect can significantly increase the classifier's sensitivity to class boundaries' precision (also referred to as "Hughes phenomenon") when the size of the training sample set can not handle the input space dimension (Jackson and Landgrebe 2001, Ho and Basu 2002). The consequence is that the classification accuracy first grows and then declines as the number of spectral channels increases while training samples are kept the same. Increasing the number of training sampling units, when it is possible, can overcome the "Hughes phenomenon", but then it incurs another risk which is that of over-fitting. More practically also, the computational cost of classification is generally a non-linearly increasing function of the feature space dimension. All these reasons highlight the necessity for a better organization of the information initially diluted in the original high dimensional input space, and several extraction/selection techniques, such as discriminant analysis, have been commonly used to perform a particular type of feature extraction (Landgrebe 2005).

In this study we evaluate the relative usefulness of high spectral and temporal resolutions of MODIS imagery data for land cover classification. Our work is a methodological study to answer the fundamental question raised by the proliferation of hyperspectral and multitemporal satellite imagery: Regarding land cover classification, what is the informational content inherent to reflectance measured in different wavelengths and at different times? We also address the corollary question: Within the same context and based on our particular data set, how to combine in a certain effective way, spectral channels with acquisition dates?, where "effective way" is meant for optimizing the trade-off between land cover differentiation and feature space dimension. In the course, we were naturally led to

evaluate the beneficence expected from inserting temporal diversity in a multispectral based classification framework.

To tackle these issues and to answer these questions, we followed a rigorous process in two main steps: 1) temporal and spectral data prioritization; and 2) use of land cover classification performances to assess the informational increment conveyed by additional temporal or spectral measures. We resort to a statistical separability analysis based on the Mahalanobis median distance to rank temporal and spectral features according to their individual and joint ability at separating land cover classes. Specific use of Mahalanobis distance was not crucial, and we could have defined the input features' prioritization based on classification results directly. However, as we wanted to keep the organization procedure as independent as possible from the classifier choice, simple arithmetic and geometric arguments led us to choose the median Mahalanobis distance as our ordering criterion.

To measure the capacity of each retained features combination at differentiating land cover classes, we used overall classification accuracy as an index to evaluate the relative information gained from each spectral or temporal dimension adjunction. As we deal with a finite data set, it is essential to minimize the influence of the searching space dimension onto the classifier performance, so that results only depend on the amount of added information. In other words, the benefits resulting from adding a new dimension to the feature space must not be annihilated by the classifier performance decrease due to dimension increase. Structurally, Support Vector Machines (SVM) (Vapnik 1998) are kernel based supervised classifiers whose performances are quite unaffected by oversized dimensions of input feature spaces (Huang *et al.* 2002, Zhu and Blumberg 2002, Marçal *et al.* 2005, Pal and Mather 2005, Mercier and Lennon 2003, Watanachaturaporn *et al.* 2004, Watanachaturaporn *et al.* 2005, Gonçalves *et al.* 2005). Inherent computational cost notwithstanding, this is the unique reason that motivated our choice of SVM for the classification task. Like this, we claim that the best classification performance should naturally be obtained with the full rank data set, and we will quote experimental results that show that indeed, the misclassification rate is constantly decreasing with the input feature space dimension.

3.2. Study area and experimental data set

3.2.1. Study area

We focused our experiences in the Portuguese mainland territory. The country has an area of approximately 89 000 Km^2 and ground altitudes raging from 0 to 2000m above sea level. Portugal is in a transition zone featuring diverse landscapes representing both Mediterranean

and Atlantic climate environments. According to Caetano *et al.* (2005), in 2000 the agricultural areas occupy 48% of the country, closely followed by forest (38%) and semi-natural areas (9%), while artificial surfaces comprise only 3% of the total land cover.

3.2.2. Land cover classes

The nine land cover classes of the developed nomenclature were defined through the Land Cover Classification System (LCCS) from Food and Agriculture Organization (FAO) (Di Gregorio and Jansen 2000) (Table 3.1). The rationale behind the development of the nomenclature was three-fold: (1) a nomenclature that is well adapted to the type of landscapes existent in regions with characteristics similar to the Portuguese mainland, (2) a nomenclature that is compatible with established ones (e.g. CORINE Land cover, Global Land Cover and the International Geosphere-Biosphere Programme nomenclatures) in order to turn possible the comparison between our results and others using different nomenclatures, and (3) a nomenclature that matches the spatial resolution of used satellite imagery. In addition, its development attended to the suggestion endorsed by Loveland *et al.* (2000), i.e. each vegetated land cover class represent relatively homogeneous land cover characteristics (e.g. similar floristic and physiognomic characteristics), that exhibit distinctive phenology (i.e. onset, peak and seasonal duration of greenness) and have comparable levels of relative primary production.

Table 3.1. Land cover classes description, label and respective number of collected samples.

Class	Label	Description	Sample size
Artificial Areas	AA	Built-up areas and transport network. Non-linear areas of vegetation and bare soils are exceptional.	47
Rainfed Crops	RC	Agricultural areas that are not artificially irrigated and consequently do not present vigorous vegetation during summer.	59
Irrigated Crops	IC	Agricultural areas irrigated artificially and periodically (during summer-time) that only present vegetation during the summer.	52
Broadleaved Forest	BF	Wooded areas where broadleaved species predominate.	40
Needleleaved Forest	NF	Wooded areas where coniferous species predominate.	55
Natural Grassland	NG	Natural areas with herbaceous vegetation, namely gramineous species.	16
Shrubland	S	Natural areas dominated by bushes and other shrubby vegetation that is characteristic of Mediterranean climate.	18
Barren	B	Area of limited ability to support life, namely thin soil, sand, or rocks.	27
Water Bodies	WB	Natural or artificial stretches of water.	40

3.2.3. Earth Observation data

Our study relies on the data acquired by the Moderate Resolution Imaging Spectroradiometer (MODIS), an instrument on board Terra and Aqua satellites from National Aeronautics and Space Administration (NASA). Terra MODIS and Aqua MODIS take between one and two days to cover the entire Earth's surface, with a complete 16-day repeat cycle. Both sensors acquire data in 36 spectral bands, or groups of wavelengths, and their spatial resolution (pixel size at nadir) is 250 m for channels 1 and 2 (0.6 µm - 0.9 µm), 500 m for channels 3 to 7 (0.4 µm - 2.1 µm) and 1000 m for channels 8 to 36 (0.4 µm - 14.4 µm). These channels are calibrated on orbit by a solar diffuser (SD) and a solar diffuser stability monitor (SDSM) system, which convert the Earth surface radiance to radiometrically and geo-located calibrated products for each band (Xiong *et al*. 2003). Although recent evaluations have reported a geo-location error of 113 m at nadir (Knight *et al.* 2006), official technical specifications warrant 50 m geo-location accuracy (Wolfe *et al.* 2002). In both cases, the geo-location is less than half a pixel dimension, and hence acceptable for our multitemporal analysis.

The data acquired by the MODIS sensor is used to generate multiple products at different pre-process stages. In this study we used the MOD09A1 product, a weekly composite of surface reflectance images, freely available from MODIS Data Product *website* (http://modis.gsfc.nasa.gov). This specific product is an estimate of the surface spectral reflectance imaged at a nominal spatial resolution of 500 m for the first seven bands as it would have been measured at ground level if there were no atmospheric scattering or absorption. The applied correction scheme compensates for the effects of gaseous and aerosol scattering and absorption, for adjacency effects caused by variation of land cover, for Bidirectional Reflectance Distribution Function (BRDF), for coupling effects, and for contamination by thin cirrus (Vermote and Vermeulen 1999). We considered a set of 43 MOD09A1 images for each one of the seven first spectral bands, covering a full year observation period, from February 2000 to January 2001. In addition, from spectral bands B2 (near infrared channel - ρ_{nir}), B1 (red channel - ρ_{red}) and B3 (blue channel - ρ_{blue}), two vegetation indices were also calculated for each date and used as additional band information, namely the *Normalized Difference Vegetation Index* (NDVI):

$$NDVI = (\rho_{nir} - \rho_{red})/(\rho_{nir} + \rho_{red}) \qquad (3.1)$$

and the *Enhanced Vegetation Index* (EVI):

$$EVI = 2.5\frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + 6\times\rho_{red} - 7.5\rho_{blue} + 1} \quad (3.2)$$

This way, each sampling unit is a time series of 43 weekly observation points lying in the nine dimensional space $[0,1]^7$ x $[-1,1]^2$ corresponding, respectively, to the reflectances and to the vegetation indices of a 500 m-by-500 m square area.

3.2.4. Sampling procedure

For each of the nine land cover classes, representative ground reference data were collected from Portuguese mainland territory along year 2000 and used for training and testing purpose of our methodology. Homogeneous sample units, each corresponding to a MODIS pixel area, were selected for each land cover class using the CORINE Land Cover map for 2000 (CLC2000) (Painho and Caetano 2006) as strata. High spatial resolution Earth Observation (EO) data, namely Landsat ETM+ images acquired in 2000 and orthorectified colour infrared aerial photographies from 1995, each covering the whole territory, were used as the base data source to recheck reference land cover classification of sample units. Reference information for each sample unit was derived by the method described by DeFries *et al*. (1998), i.e. by visual interpretation of high resolution EO data overlaid with a co-registered 500 m fishnet corresponding to the MODIS data grid. At such a coarse resolution, land cover homogeneity is hardly guaranteed by random sampling because most pixels are likely to contain features from two or more distinct classes. Thus, we deterministically retained only pixels that are representative of pure land cover class occupation. As our intention is not to produce a thematic map based on a new operative classification procedure, it was fundamental to consider empirical classes with typical spectral and temporal specificities that are statistically sound. In the collect process, we endeavored at spreading as much as possible each class samples over the mainland territory, in order to account for possible regional differences, and to prevent geographic correlation due to adjacent pixels (Hammond and Verbyla 1996). Finally, as our objective is to assess the pertinence of intra-annual temporal information for land cover characterization, each sampling unit was selected within geographical sites that did not undergo a drastic land cover change during the study period.

Although the appropriate number of sample units per class that are needed to train a supervised classifier remains open to debate, certain hints or simple heuristics relate the strict minimum to (*i*) the used algorithm itself, to (*ii*) the number of input variables, to (*iii*) the sample selection method, and to (*iv*) the size and spatial variability of the study area (Huang *et al.* 2002, Jensen 1996, Mather 2004, Foody and Mathur 2006, Foody *et al.* 2006).

However, Ho and Basu (2002) asserted that certain classification problems have nonzero Bayes error, no matter the sample size or the feature space dimension.

3.3. Data analysis

3.3.1. Pre-process of Earth Observation data

Despite all the default corrections performed on MODIS products, data pre-processing is mandatory to avoid biases due to time-specific contaminations. We were led to develop some additional treatment for cloud detection and removal, and for the intra-annual calibration of the spectral profiles of land cover classes. Firstly, we choose to base our study on weekly composites of the spectral reflectances and vegetation indices, rather than on daily values. Compositing is a process that selects for each pixel-size area of an image the "best" pixel among all available within a given period of time (in most cases, optimality is meant for "cloud-free" situations). Clearly, this selection implicitly assumes land cover to be locally stationary in time, which in turn restricts the window extent to a sensibly small period of time. Our study is particularly sensitive to this limit as it entirely relies on the land cover specific intra-annual variations of reflectance. Therefore, we are forced to a reasonably short compositing window that will not smooth out the characteristic non-stationarities of the time series, but sometimes not sufficiently large to remove all clouds presence. Consequently, measured reflectances may still locally exhibit abnormally high values, similar in kind to a shot noise superimposed to the relevant signal (Figure 3.1).

According to Van Dijk *et al.* (1987), a time-sliding windowed median filter outperforms other smoothing techniques, such as linear filters or polynomial fits, at removing aberrant measures from the time series. This approach was successfully employed by Duchemin *et al.* (1999) on NDVI time-series to monitor the phenology of various forest ecosystems, and that is the technique we also retain to denoise our data streams. Each sample of the original time series $x[k]$ is replaced by the Q-th percentile of the short time windowed sub-series $x[n] \cdot W[n-k]$ centered on the sliding index k. The Q-th percentile is a free parameter that can be tuned to adapt to the shot noise amplitude, and that we empirically fixed to 40%. The criterion that guides our choice of both the Q-th percentile and the window length is purely heuristic. We did extensive trials with different values of these parameters and compared the resulting responses in all classes. We then empirically set them to the values, which smoothed out local oscillations and aberrant measures while preserving the slow varying profiles (corresponding to a yearly period). Notice though, that because we use a narrow window that does not comprise many samples (window width does not exceed six weeks),

there is no dramatic difference for Q's lying in the interval [30%-50%]. We then applied the same short-time median filter with the same set of parameters to all traces. Regarding the vegetation indices, their non-linear calculation does not commute with the median filtering. Yet, comparing the results when permuting the order of operations does not reveal drastic changes. In practice, we did apply the median filter to EVI and NDVI computed from the original spectral bands.

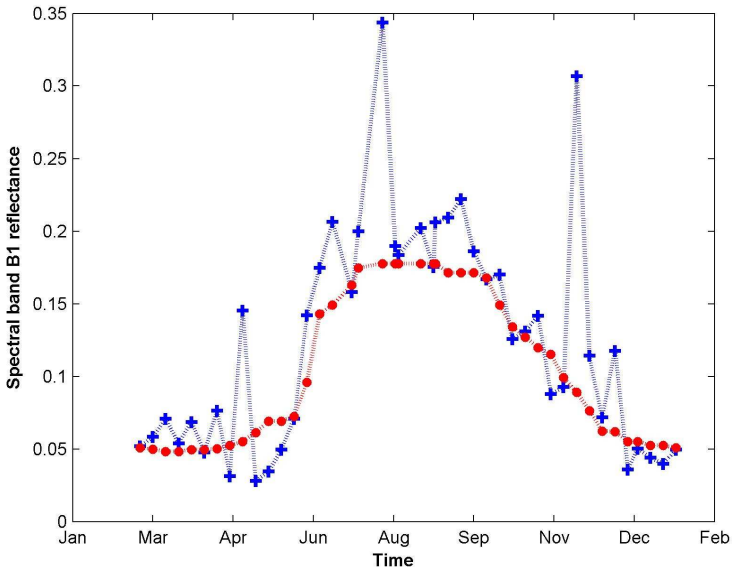

Figure 3.1. Typical time series corresponding to the intra-annual evolution of reflectance in MOD09A1 spectral band B1 (red wavelength) for a pixel area of the Rainfed Crop (RC) class. Crosses correspond to the raw observed data. Dots show the output of a time-varying windowed median filter (we used a gaussian window with a width corresponding to a six weeks duration).

From the example of Figure 3.1, it is clear how the reflectance bursts due to clouds' presence disappear from the non linear filter output, without the underlying trend being altered. Furthermore, the low frequency content of the resulting response can easily withstand down sampling, and in the sequel, we shorten the original weekly time series to 12 points, retaining only one measure per month.

3.3.2. Statistical ordering of input features for classification

Innovative approaches are constantly proposed to tackle the challenging issue of extracting the most condensed and the most pertinent information from a highly redundant feature space (Landgrebe 2002, 2005). In many applications, like in automatic land cover classification from satellite imagery databases, the difficulty gets worse when the number of

available training samples – too small to cope with the high dimensionality of the problem – hampers effective classification results (Ho and Basu 2002, Guyon and Elissee 2003).

Having in mind the Bayes error as a theoretical bound, but difficult to attain in practice, a large number of metrics have been devised to measure the ability of the features, taken individually or in conjunction, at discerning the classes. Among all existing approaches based either on probability distances (e.g. Bayesian distance), entropy measures (e.g. the Kullback-Leibler distance), correlation, or geometry (e.g. the Fisher ratio), we chose a technique based on "statistical separability indices" (Bruzzone *et al.* 1995). Many statistical distances such as Bhattacharyya, Jefries-Matusita, Transformed Divergence, or Mahalanobis have been reported as relevant criteria for selecting the adequate feature subspace in a two-classes classification problem (Lee and Landgrebe 1993, Bruzzone *et al.* 1995, Haertel and Landgrebe 1999, Jimenez and Landgrebe 1999, Herold *et al.* 2003, Marçal *et al.* 2005, Watanachaturaporn *et al.* 2005). Naturally, these measures are strongly correlated (Singh 2003). Now, considering a multi-class classification problem, it is not so clear how to combine these two-classes' distances to get a global dispersion measure. A straightforward way is to compute the average value over all pairs of classes' distances (Bruzzone *et al.* 1995, Watanachaturaporn *et al.* 2005). Unfortunately this approach does not account for the distances heterogeneity, and a distance between two classes that is notably larger than all the others, can leverage the mean distance and mask the overall classification difficulty. Conversely, a MinMax algorithm selecting the feature subspace that yields the best separation between the two closest classes (Lee and Landgrebe 1993, Haertel and Landgrebe 1999, Jimenez and Landgrebe 1999, Herold *et al.* 2003), does not necessarily guarantee an optimal separability among the other classes.

Here we propose to use the median of the Mahalanobis inter-class distances as our sorting criterion to ordinate the temporal and the spectral features in our land cover classification framework. The Mahalanobis distance between two points belonging to a *p*-dimensional space is defined as follows (Johnson and Wichern 1998):

$$d_s = \sqrt{(x-y)^t S^{-1}(x-y)} \qquad (3.3)$$

where *S* is the covariance matrix.

Considering a particular feature subspace *F* – in our case a combination of spectral bands and/or of dates – we first calculate all pair wise Mahalanobis distances that separate each sample of a given class *n*, from all samples of the remaining classes. Doing this for all classes $n=1, ..., 9$, we then compute the median value *D(F)* of all the resulting pairwise

distances, and use this as a measure of the separation strength for the considered feature subspace. The median (Mahalanobis) distance is a global measure that groups all samples of a same cluster into a unique distance to the remaining population. Calculating this distance for all the nine classes individually corresponds to a "one-versus-the-rest" strategy, commonly used to generalize a two-class problem to a multi-class situation.

On the other hand, choosing the median rather than the mean value blurs the dominant effect of a few atypical large distances, as the median equitably balances the two halves of a population. Notice also that in the process we do not take into account the distances between sampling units of a same class, meaning that only through the variance matrix does the own class dispersion intervene in this sorting process. Eventually, the various feature spaces are ranked according to their corresponding median inter-class distance: the larger this distance, the more spread out and supposedly separable are the classes when represented in this feature space.

We apply this procedure to our multispectral images time series, in order to determine the p (bands) x q (dates) dimensional feature subspace $F_{p,q}$ that maximizes the Mahalanobis median distance $D(F_{p,q})$. We start considering spectral bands and dates separately, and find the feature space $F_{1,1,}$ that among all possible combinations of a single band with a single date maximizes the Mahalanobis median distance. We call the corresponding input feature $m_{i1,j1}$, i.e. the measure from channel B_{i1} (i_1=1,...,9), at month t_{j1} (j_1=1,...,12), and $D(m_{i1,j1})$ the associated distance.

Finding the optimal combination of dates

Once, $m_{i1,j1}$ is fixed, we form all possible combinations $F_{1,2}=\{m_{i1,j1}, m_{i1,j2}\}$ with $j_2 \neq j_1$, and calculate for each one the median Mahalanobis distance $D(F_{1,2})$. The maximum distance designates the best class dispersant representation plane that is spanned by only one spectral reflectance (or vegetation index) B_{i1} measured at two different dates t_{j1} and t_{j2}. Similarly for the remaining dates, each combinatorial subset $F_{1,q}=\{m_{i1,j1}, m_{i1,j2}, ..., m_{i1,jq}\}$ for q=1,...12, defines the most judicious way of arranging the spectral measure from q different dates of the year.

Finding the optimal combination of spectral bands and vegetation indices

Following the same bottom-up algorithm, and considering the best single date t_{j1}, for p=1,...,9, we determine the combinatorial subsets $F_{p,1}=\{m_{i1,j1}, m_{i2,j1}, ..., m_{ip,j1}\}$ with maximum Mahalanobis median distance $D(F_{p,1})$. These indicate the most relevant sequential order for including spectral reflectances and vegetation indices B_i's in the classification input space. In

particular, $F_{9,1}$ designates the period of the year that *a priori* offers the best full spectral conditions for land cover classification in Portugal.

As a general remark, it is naturally expected that the Mahalanobis median distance increases with the dimension of the feature spaces $F_{p,q}$, and ultimately, any arbitrary ordering will reach the same distance. Our concern here was to rank the feature inputs so as to obtain the fastest convergence towards the best expectable classification results. The next step is to show the role of this arrangement on classification performance, and to appraise the relative gain obtained at each supplementary temporal or spectral feature inclusion. Before that, the following section briefly describes Support Vector Machines, a supervised classifier commonly considered as immune to a performance decrease due to dimension increase.

3.3.3. Land cover classification

Support Vector Machines (SVM) are a new generation of supervised learning systems based on recent advances in statistical learning theory (Cristianini and Shawe-Taylor 2000). Pioneered by the work on learning strategy by Vapnik and collaborators (Boser *et al.* 1992, Vapnik 1998), they have rapidly and successfully been applied to numerous real-world classification problems.

Conceptually, SVM rationale is that a classification problem that does not have a satisfying solution in its own observation space may have one simple and efficient in a more complicated representative system. As so, SVM use a hypothesis space of linear indicator functions to draw classification in a high dimensional (possibly infinite) feature space, image of the observation space by a non-linear mapping Φ. For the sake of simplicity we restrict this introductory study to the case of a two classes' classification. Generalization to multiple classes' problems is straightforward using, for instance, a one-versus-the-rest strategy.

Consider a binary classification-learning task and the following data training set:

$$S\{(\mathbf{x}_1, y_1),...,(\mathbf{x}_N, y_N)\}, \mathbf{x}_i \in \mathbf{X} \subseteq \Re^{n_x}, y_i \in \{-1,+1\} \qquad (3.4)$$

with samples $(\mathbf{x}_i, y_i)$ drawn i.i.d. according to some unknown but fixed probability distribution $P(\mathbf{x}, y)$. The non-linear mapping Φ transforms the n_x-dimensional input space $\mathbf{X}$ into the feature space $\mathbf{F}$, and let the set of hypotheses be linear hypothesis functions of the type:

$$h(\mathbf{x}) = \sum_{i=1}^{N} \alpha_i y_i \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}) \rangle + b \qquad (3.5)$$

One remarkable fact about this hypothesis function (written here in its dual form) is that it implies the data only through their inner products in the feature space. Therefore, if Φ is properly chosen so that it obeys the so-called *kernel* condition,

$$\langle \Phi(\mathbf{x}), \Phi(\mathbf{z}) \rangle = K(\mathbf{x}, \mathbf{z}), \tag{3.6}$$

we do not even need to know the underlying feature map Φ to be able to learn in the feature space.

Assuming so, the problem is then to determine the coefficients $\{\alpha_i, i = 1,..., N\}$ that minimize the classification error on unseen samples. Ideally, the best classifier should minimize the expected value of the *loss* or *risk*:

$$R(h) = \mathrm{E}_{P(\mathbf{x},y)}\{L(y, h(\mathbf{x}))\} = \int L(y, h(\mathbf{x}))dP(\mathbf{x}, y), \tag{3.7}$$

where the loss function $L(y, h(\mathbf{x})) \geq 0$, penalizes the deviations. In practice however, the joint probability function $P(\mathbf{x}, y)$ is unknown, and the expected loss (also called *generalization loss*) is approximated by the *empirical classification error* based on the available information (i.e. the training set). Then, the empirical risk functional reads:

$$R_{emp}(h) = \sum_{i=1}^{N} L(y_i, h(\mathbf{x}_i)) . \tag{3.8}$$

Often though, finding the hypothesis h^* that minimizes this empirical risk leads to an ill-posed problem that results in the well-known phenomenon described as *over-fitting* in the literature (i.e. the selected hypothesis is too complex). One way to avoid over-fitting tolerating noise and outliers is to restrict the complexity of the hypothesis function, by introducing a regularization term (Vapnik 1982). This regularization term is closely related to the notion of margin, another important concept for SVM that reflects the sensitivity and tolerance of the classifier to the samples $\{\mathbf{x}_i, i = 1,..., N\}$ that stand "close" to the (non-linear) separator. Those points are called support vectors, and the solution that minimizes the regularized empirical risk function is referred to as a soft margin SVM (in contrast to a hard margin SVM that systematically zeroes the empirical loss). Solution to this quadratic optimization problem is classically obtained by Lagrangian theory and comes out to find the Lagrangian multipliers α^* that maximize the following quantity:

$$W(\alpha) = \sum_{i=0}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} y_i y_j \alpha_i \alpha_j (K(\mathbf{x}_i, \mathbf{x}_j) + \frac{1}{C} \delta_{ij}) , \tag{3.9}$$

subject to $\sum_{i=1}^{N} y_i \alpha_i = 0$ and $\alpha_i \geq 0$, i=1,…N.

C is the regularization parameter that bounds the Lagrangian multipliers (i.e. the weights associated to the support vectors) controlling this way the capacity of the hypothesis function class. There exists an alternative way for controlling the hypothesis function capacity, which amounts to fix the proportion of training samples that will lie between the boundary and the margin hyperplane. In SVM literature, this regularization strategy is often referred to as the ν-parameterization (Scholkopf *et al*. 2000), and we used this in our study.

Finally, a Radially Basis Function (RBF) kernel, $K(\mathbf{x},\mathbf{z}) = e^{-\frac{\|\mathbf{x}-\mathbf{z}\|^2}{\sigma^2}}$, with user pre-defined sample variance σ^2 is used. We chose Gaussian radial basis functions simply because they yield classification results that outperformed those obtained with all the polynomial kernels we tested on our data set. In that case, the corresponding feature space is a Hilbert space of infinite dimension, which is commonly considered as the most versatile searching space for linear discrimination (Suykens *et al.* 2002, Seeger 2004).

Very little guidance exists in the literature on which criteria to use for selecting the kernel-specific parameters of SVM. In our case, we resort to a systematic and computationally expensive scan to determine the set of hyper-parameters values minimizing the class loss. This empirical optimization needs to be done for each input space we use. Moreover, all experimental results that we report in this work derive from cross validation, a standard procedure used in longitudinal data analyses, such as land cover classification (Huang *et al.*, 2002, Foody *et al*. 2006). Given a supervised classifier and a unique training sample set, cross validation allows for assigning a confidence interval to the expected loss estimate. After splitting the data set into K folds, each fold is successively used for testing purpose while the classifier is trained on the remaining data. Thereby, averaging the resulting K empirical classification errors yields an estimate of the expected loss, and the standard deviation plays the role of a confidence interval. In our study, we fixed to five the number of cross validation folds by taking into account the sample size per land cover class. This repartition yielded the best compromise between bias and variance of classification loss estimate (also referred to as "bias-variance trade-off"). Increasing the number of folds would minimize the variance of our estimation, but increase the bias, and conversely when decreasing the number of folds.

From a practical perspective, we used SVM software of Spider Version 1.6, an object orientated environment for machine learning in MATLAB®, and for a more pragmatic presentation of SVM, see, e.g. Huang *et al.* (2002) and Foody *et al*. (2006).

3.4. Results and discussion

3.4.1. Temporal and spectral data prioritization

We applied the pre-processing technique of section 3.3.1 to the multitemporal and multispectral sampling units collected as described in section 3.2.4, and considered this output as our working data. The Mahalanobis median distances described in section 3.3.2 are calculated on this working data set in order to determine the arrangements of dates and spectral channels that guarantee the maximum spacing between the nine land cover classes prior to the classification.

The best singleton $F_{1,1}=\{m_{i1,j1}\}$ coincides with the EVI channel measured in August, indicating that in summer the land cover classes stand further apart along the EVI feature than along any others ($D(F_{1,1})$ in Table 3.2). In agreement with conclusions reported in Borak and Strahler (1999), vegetation indices integrate various discriminating characteristics, as they form non linear combinations of reflectances in different spectral bands. Moreover, according to Xiao *et al.* (2003) and to Boles *et al.* (2004), EVI is more correlated to vegetation phenology than NDVI because it is less sensitive to atmospheric conditions, soil background and it does not saturate so easily with closed vegetation canopies. Regarding the second and the third components of this spectral channels' arrangement ($m_{i2,j1}$ and $m_{i3,j1}$), it is not so surprising to find reflectances in spectral bands B7 (short-wavelength infrared) and B4 (green wavelength), respectively, as these later are two channels located at the opposite ends of the spectrum, thus weakly correlated, and none of them entering the definition of EVI.

Now, regarding the best date t_{j1}, optimality holds from a global viewpoint only. Since summer time is usually a drought season in Portugal, it is most likely that among all the classes, those who present vegetation dominance will produce very similar spectral reflectances during this period. This is clearly illustrated in Figure 3.2, where the averaged EVI profiles of Barren (B), Natural Grassland (NG) and Rainfed Crops (RC) classes, practically mix up during the summer time, while they distinctly differ in winter. It is precisely this complementary information that is the stake of feature inputs' arrangements. So, pursuing the bottom-up approach described in section 3.3.2, we end up with the best full dates' ordering disclosed in Table 3.2. As expected, after August, a winter month

(t_{j2}=February) enters in second position of the arrangement, providing with the complementary information that is most able at increasing the median inter-class distance. Finally, although it was individually the second most classes-dispersant month, because July is strongly correlated with August, it only appears in the ninth position of the ranking.

Table 3.2. Mahalanobis median distances $D(F_{p,q})$ between the nine land cover classes. $F_{p,q}$ corresponds to the feature subspace composed of the p channels $\{B_{i1},\ldots, B_{ip}\}$ measured at the q dates $\{t_{j1}, \ldots, t_{jq}\}$.

		B_{i1}	B_{i2}	B_{i3}	B_{i4}	B_{i5}	B_{i6}	B_{i7}	B_{i8}	B_{i9}
		EVI	**B7**	**B4**	**B3**	**B1**	**NDVI**	**B5**	**B2**	**B6**
t_{j1}	**Aug.**	1,25	1,80	2,23	2,77	3,07	3,32	3,57	3,91	4,16
t_{j2}	**Feb.**	1,68	2,67	3,26	3,87	4,32	4,70	5,12	5,57	5,90
t_{j3}	**Dec.**	2,27	3,34	4,15	4,74	5,31	5,73	6,31	6,74	7,16
t_{j4}	**May**	2,63	3,90	4,80	5,52	6,15	6,72	7,23	7,79	8,22
t_{j5}	**Jan.**	3,08	4,15	5,28	6,15	6,83	7,49	8,15	8,73	9,25
t_{j6}	**Nov.**	3,27	4,41	5,69	6,67	7,47	8,33	9,02	9,55	10,20
t_{j7}	**Jun.**	3,37	4,65	6,07	7,16	8,00	8,84	9,55	10,22	10,87
t_{j8}	**Oct.**	3,57	5,06	6,44	7,55	8,44	9,38	10,18	10,90	11,68
t_{j9}	**Jul.**	3,69	5,45	6,86	8,01	8,96	9,91	10,93	11,59	12,38
t_{j10}	**Apr.**	3,76	5,68	7,24	8,46	9,53	10,53	11,53	12,18	13,05
t_{j11}	**Mar.**	3,93	5,87	7,66	8,96	10,10	11,08	11,94	12,77	13,66
t_{j12}	**Sept.**	3,99	6,02	8,00	9,38	10,55	11,50	12,44	13,38	14,25

Notice that we could have followed a slightly different prioritization scheme, by selecting after the most informative single channel, the next one from any of the remaining 107 channels. This particular ordering would have then led to the "best" possible combination of n ($1 \leq n \leq 108$) time-spectral measures among all observations. However this approach does not permit to dissociate the gain obtained when the feature space dimension is incremented along either one of the time or of the spectral direction, while the other is maintained fixed. As our primary objective was to appraise time and frequency contributions independently, we intentionally opted for a strategy that does preserve the marginal properties of Table 3.2.

3.4.2. Land cover classification analysis

We considered the arrangement of bands and dates obtained at previous stage, as the most effective scheduling for exploiting reflectance information in land cover classification. Our objective here was two-fold: On the one hand, we wanted to appraise the relative contribution of each spectral or temporal component to land cover classification. On the other hand, we seek at correlating the inter-class Mahalanobis distance growth with the evolution of land cover classification accuracy, as the amount of available information increases.

Starting with $F_{1,1}$ of Table 3.2, and following the same order for successively adding reflectance measured at different wavelengths or at different times, we trained and tested a SVM on our working data set, for each feature subspace $F_{p,q}$. Corresponding overall classification accuracies $C(F_{p,q})$ and significance of the differences between $C(F_{9,12})$ and each $C(F_{p,q})$ are reported in Table 3.3.

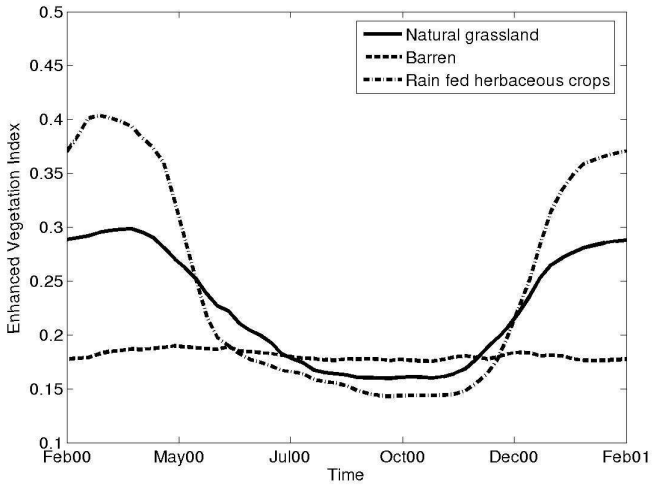

Figure 3.2. EVI time profiles obtained by averaging the sample belonging to the same land cover class. The three displayed classes, Natural Grassland (NG), Barren (B) and Rainfed Crops (RC), mix up in summer time, whereas they are perfectly distinguishable in winter.

Table 3.3. Overall classification accuracy $C(F_{p,q})$ of the nine land cover classes, obtained with an optimized SVM trained and tested on $F_{p,q}$. Each $F_{p,q}$ corresponds to the feature subspace composed of the p channels $\{B_{i1},\ldots, B_{ip}\}$ measured at the q dates $\{t_{j1}, \ldots, t_{jq}\}$.

		B_{i1}	B_{i2}	B_{i3}	B_{i4}	B_{i5}	B_{i6}	B_{i7}	B_{i8}	B_{i9}
		EVI	**B7**	**B4**	**B3**	**B1**	**NDVI**	**B5**	**B2**	**B6**
t_{j1}	**Aug.**	0,14**	0,69**	0,78**	0,85*	0,88	0,87	0,88	0,89	0,89
t_{j2}	**Feb.**	0,74**	0,79**	0,84**	0,86*	0,86*	0,87	0,88	0,88	0,88
t_{j3}	**Dec.**	0,79**	0,83**	0,86*	0,87	0,86*	0,88	0,89	0,89	0,89
t_{j4}	**May**	0,79**	0,84**	0,88	0,88	0,88	0,88	0,89	0,89	0,89
t_{j5}	**Jan.**	0,79**	0,84**	0,86*	0,88	0,88	0,88	0,89	0,89	0,89
t_{j6}	**Nov.**	0,79**	0,84**	0,88	0,88	0,88	0,88	0,89	0,89	0,89
t_{j7}	**Jun.**	0,80**	0,85*	0,88	0,89	0,89	0,89	0,89	0,90	0,90
t_{j8}	**Oct.**	0,81**	0,85*	0,88	0,89	0,89	0,89	0,89	0,90	0,90
t_{j9}	**Jul.**	0,83**	0,86*	0,88	0,89	0,90	0,89	0,89	0,90	0,90
t_{j10}	**Apr.**	0,82**	0,86*	0,88	0,89	0,90	0,89	0,88	0,90	0,90
t_{j11}	**Mar.**	0,81**	0,85*	0,89	0,90	0,88	0,89	0,88	0,90	0,90
t_{j12}	**Sept.**	0,82**	0,85*	0,89	0,89	0,89	0,89	0,89	0,89	0,90

*Significant at 5%, ** Very significant at 1%

As a preliminary remark, let us comment on the classification accuracy obtained with $F_{1,1}$. This baffling low value (as compared to the others) undoubtedly comes from the inability of SVM at transforming a complex decision boundary in a one-dimensional original input space into a linear frontier in a higher dimensional feature space (Huang *et al.* 2002). To support our claim and to verify that this abnormally poor performance is not due to insufficient discriminating data, we trained and tested, under the same conditions, a K-Nearest Neighbors (KNN) algorithm – a commonly used supervised classifier. As expected, we obtained with $F_{1,1}$, a significantly larger overall classification accuracy of 56,4%.

Previous studies never reported on this SVM limitation because classifications with large numbers of spectral features were systematically considered to demonstrate superiority of SVM over other classifiers, e.g. Shah *et al.* (2003), Marçal *et al.* (2005), and Pal and Mather (2005).

We now analyze in detail the evolution of the classification accuracy when increasing the feature space dimension. In the first place, it is important to notice that classification globally improves at each supplementary information inclusion; so that the best performance (about 90% matches) is attained combining all spectral bands and dates. However, and more importantly, the classification accuracy reaches a very acceptable rate with quite a few number of feature inputs. Indeed, with only 4 spectral measures {EVI, B7, B4, B3} taken at 4 dates {Aug., Feb., Dec., May}, the classification $C(F_{4,4})$ scores an 88% match that is slightly inferior to the best performance, but not significantly so (Table 3.3). Then, any additional spectral or temporal feature does not drastically improve the performances that end stabilizing around 90%.

To some extent these results are contradictory with the ones reported in Pal and Mather (2005) that related a more constantly increasing classification rate with the feature space dimension. The DAIS sensor used in their study provides with reflectance in 72 spectral bands spanning a total bandwidth between 400 nm and 2500 nm, approximately equal to the MODIS spectral range (400 nm – 2100 nm) that we use in the present work. Integrating successively the channels into the feature space of a SVM classifier, the classification accuracy starts at 65% with 5 bands and only stabilizes around 93% beyond a 50 dimensional feature space! We can evoke two reasons to explain this divergence: one is simply instrumental and the other is methodological. It is very likely that because of spectral correlation, the 72 bands of the DAIS sensor (versus 7 for the MODIS sensor) present a high degree of redundancy, which consequently diminishes the individual contribution of each reflectance measurement to land cover characterization. If that is the case, much more DAIS

channels (50) are needed to bring in the equivalent amount of information corresponding to only few MODIS channels (4). From a methodological viewpoint, the prime objective of our spectral bands arrangement was to obtain the fastest convergence towards the best expectable classification results. Besides, resorting to vegetation indices that inherently encompass two (NDVI) or three (EVI) different spectral bands, is by itself a first attempt to optimize spectral information organization. In this respect, the resulting effect of an adequate ordering on the rate of convergence of the classification accuracy shows quite convincing from our experiments.

We now proceed analyzing the information content of multitemporal versus multispectral measures. It is clear from the classification rates obtained with $F_{9,1}$ and with $F_{1,12}$, respectively, that spectral diversity is a richer source of information than time variety. Reflectance in all spectral bands measured at a single date (should this date be the most discriminatory of the year), is almost enough to achieve the best classification performances (89% accuracy), whereas a full year measurement of the Enhanced Vegetation Index does not suffice to compensate for the missing spectral information (82% accuracy). Important also is the fact that there is a significant difference between $C(F_{1,12})$ and $C(F_{9,12})$, while $F_{9,12}$ did not generate significantly better classification results than $F_{9,1}$ (Table 3.3). Basically, once the feature space integrates the five most discriminative spectral bands, multitemporality turns practically useless. In contrast, even with the twelve dates, there is always substantial interest in considering no less than three spectral channels. These conclusions are quite corroborating with earlier studies, which all claim that spectral data are preferable to multitemporal data for their classification skills. For instance, Borak and Strahler (1999) maintain that all MODIS spectral bands are effectively necessary for land cover classification, whereas the temporal domain is not so relevant. As for us, we have a more finely-shaded opinion. Based on the results of Table 3.3, it deems important to qualify the effectiveness – or not - of multitemporal measures to meliorate classification precision. In particular, it is indubitable that when a few number of spectral bands is simultaneously used (up to four in our experiences) the inclusion of several measurement dates significantly increases the classification match. That is how the classification rate rises from 69% with $F_{2,1}$ up to 85% with $F_{2,12}$, a 15 points gain uniquely due to additional time information from duplicated measures of EVI and B7 reflectance along the year. In the same direction, a set of 12 monthly measures of the triplet {EVI, B7, B4} ($F_{3,12}$) provides with the same discriminatory information as a single date measure of all spectral channels ($F_{9,1}$). To a very large extent, our conclusions sustain previous studies that already evidenced multitemporal data usefulness for land cover classification at regional and global scales, e.g. Loveland *et al.*

(1991), Zhu and Evans (1994), and Liu *et al.* (2003), to cite but a few. Nevertheless, it is important to state that these attempts to characterize land cover over large areas used AVHRR data, which comprise land observation in few broad spectral bands. Thus, and according to our previous outcomes, it is not surprising that multitemporal analyses of AVHRR spectral data enhanced land cover classification performances. Besides, the contribution of images time series to land cover classification was corroborated in our study, when only few MODIS spectral channels were considered.

We now would like to elaborate on our questionable decision to integrate vegetation indices directly as spectral inputs. In fact, as NDVI (respectively EVI) is a non linear combination of reflectances in two (respectively three) wavelengths, any single-date measure already encompasses a multispectral dimension. This creates an inequity between the different spectral channels, which may attenuate or exaggerate the effective influence of reflectance in a particular bandwidth, and alter the multitemporal impact on classification improvement. That is why we reproduce the exact same input features prioritization and classification evaluation, but discarding the vegetation indices. In short, the resulting channels' ordering that maximizes information concentration in both time and spectral directions is necessarily different from the previous configuration. In particular, reflectance in B3 (blue wavelength) and measured in December appears to be the most discriminative channel. But, more importantly, although the progression of the classification accuracy along the descriptive variables follows a sensibly similar pattern to that of Table 3.3, the plots of Figure 3.3 call for a couple of nuanced comments. First of all, the best scores achieved when including and when discarding the vegetation indices, are identical and equal to 90% match. This clearly shows that EVI and NDVI convey redundant information once all spectral channels and dates are considered simultaneously in a classification framework. Singularly, this is not the case for single date measurements, since then, vegetation indices seem to be a valuable source of information (76% versus 89%). But when EVI and NDVI are discarded, adding temporal data can compensate the resulting informational deficit! Noticeably also, if we compare the temporal evolution of classification accuracy in both situations (e.g. using only the best band), it becomes clear that the time feature contributes more significantly to land cover characterization when vegetation indices are discarded. However, in that case, the temporal content is never sufficient to compensate the absence of EVI (or of NDVI).

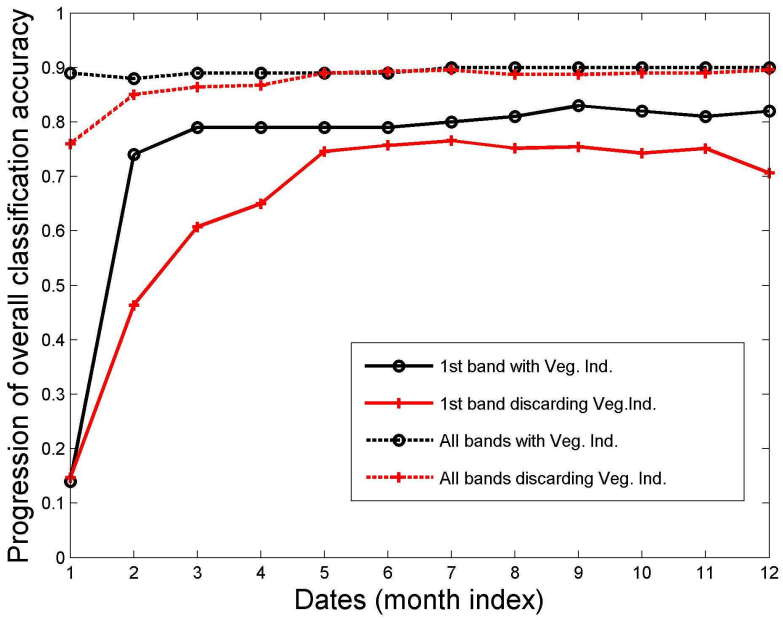

Figure 3.3. Overall classification accuracy progression when increasing the number of measurements along time. For the two situations: comprising and discarding the vegetation indices, classifications based on a single (best) channel and based on all channels are presented.

Independently of overall classification accuracy progression along spectral and temporal feature spaces increment, a closer view at the classification loss repartition (Table 3.4), shows that essentially three land cover classes are responsible for the overall mismatches (for all the other classes, both user's and producer's accuracies score above 80% and therefore are discarded from Table 3.4). Moreover, adding spectral or temporal reflectance measures as new descriptive variables to the feature space does not substantially reduce the confusion rates of those classes. As a matter of fact, once identified the systematically misclassified pixels, even when resorting to aerial imagery, it is actually rather questionable to settle the classes they belong to. All this leads us to believe that the problem originates from the theoretical definition of certain land cover classes which are so spectrally similar, that none of the temporal profiles could serve as a meaningful discriminatory parameter. A similar situation was previously reported by Hansen *et al.* (2000), which concluded that approximately half of their classification error was due to confusion between two classes only, namely the woodland and grassland classes.

The rapid convergence of the classification accuracy towards its maximum value, compared to the evolution of Mahalanobis inter-class distance under the same conditions, raises another interesting comment. In contrast to what happens with the classification rate, the inter-class distance keeps growing as the feature space dimension increases and does not

seem to attain any asymptotical value. While the median distance between classes can not straightforwardly transpose in terms of classification ability, it certainly is a good indicator of the classes' separability. Thus, as the overall distance increases, one would naturally expect the classification performance to improve as well (or at least to remain constant). Our interpretation is the following: until a certain threshold, adding descriptive variables into the feature space brings in new expanding directions along which intermingled clusters can eventually dissociate. Accordingly, the classification score jumps from 56% (KNN value) in the one dimension space $F_{1,1}$ to 88% in a sixteen dimensions space $F_{4,4}$, while the median Mahalanobis distance increases fourfold. Once the clusters are disjoint, because of the triangular inequality, any dimension adjunction will inevitably increase the distance between classes, but it will remain with no effect on outlier samples that got trapped in a misclassified cluster. The classification loss stabilizes, no matter what the feature space dimension, although the interclass-distance almost triples between $F_{4,4}$ and $F_{9,12}$. This phenomenon definitely denounces the appropriateness of a land cover typology that may not be recoverable from spectral reflectance information only, as we already mentioned it.

Table 3.4. User's and producer's accuracy corresponding to classes that mix easily: Natural Grassland (NG), Barren (B) and Shrubland (S). Values correspond to SVM classifications based on two different input feature spaces (see Table 3.2).

	$F_{4,4}$		$F_{9,12}$	
	User's accuracy	**Producer's accuracy**	**User's accuracy**	**Producer's accuracy**
NG	45,5	31,3	66,7	37,5
B	68,0	63,0	72,0	66,7
S	54,5	33,3	54,5	33,3

All the accuracy results we presented derive from the reference sample selected with the sole objective of training and testing the supervised classifier. These results are acceptable for methodological elaboration and for evaluating the effect of temporal-spectral dimensions of the feature space on land cover classification. However, they are of limited use to assess accuracy for the whole study area, because the necessary probability foundation to permit generalization from sample data to the entire population was at stake in the current study (Stehman and Czaplewski 1998). Nonetheless, our experiments convey a sensible measure of comparison between competing classifications schemes built upon different input data sets (Loveland *et al.* 2000).

3.5. Conclusion

The goal of our study was mainly to highlight the individual and combinatorial influence of the spectral and of the temporal components of remotely sensed images reflectances in land cover classification. As our aim was not to propose an operational classifier directed at thematic mapping based on the most efficient combination of reflectance inputs, we intentionally restricted our experimental framework to continental Portugal. In the course, we were led to define a distance-based topography to rank these features in terms of their relevant contribution to a discriminatory representation space. The chosen Mahalanobis median distance, although it was not shown to strictly optimize any classification criterion, deemed among all the possible arbitrary choices a reasonable indicator of the classes' dispersion and of the clusters compactness. Then, following this resulting arrangement of features inputs to sequentially train a Support Vector Machine classifier, we were able to demonstrate that the Enhanced Vegetation Index (EVI) calculated in August was the most informative combination of one spectral band with one date to characterize the land cover classes that we retained to describe the Portuguese mainland. Continuing to gradually include the remaining bands and dates, we also exposed the context dependent advantages of each new component to the classification performances, and thus proved the multitemporality assets and limitations. In this way, we showed that spectral diversity is a richer source of information than time variety. In fact, the multitemporal factor has a significant effect when coupled with combinations of few spectral bands, but it turns negligible as soon as the full spectral information is exploited. In contrast, even with a full year measurement, there is always substantial interest in considering no less than three spectral channels. As a by-product of our study, we evidenced the poor adequacy of spectral and temporal recourses at differentiating certain land cover classes. A situation often pointed out by previous investigations in diverse bioclimatic study areas, and due to baffling temporal and spectral similarities between distinct classes' phenologies.

3.6. Acknowledgements

This study was carried out in the framework of the project "LANDEO - User driven land cover characterisation for multi-scale environmental monitoring using multi-sensor earth observation data (PDCTE/MGS/49969/2003)" funded by "Programa Dinamizador das Ciências e Tecnologias para o Espaço" from "Fundação para a Ciência e Tecnologia", and from the "Announcement of Opportunity for the Utilisation of ERS and ENVISAT Data" from European Space Agency (ESA). Research by Hugo Carrão was founded by the "Fundação para a Ciência e Tecnologia" (SFRH/BD/18447/2004). This work was partially

performed while Hugo Carrão was visiting INRIA Rhône-Alpes granted with a six month INRIA scholarship.

3.7. References

Ballantine, J.A.C., Okin, G.S., Prentiss, D.E. and Roberts, D.A. (2005), Mapping North African landforms using continental scale unmixing of MODIS imagery. *Remote Sensing of Environment*, 97, 470-483.

Boles, S., Xiao, X., Zhang, Q., Munkhutya, S., Liu, J. and Ojima, D.S. (2004), Land cover characterization of Temperate East Asia, using multi-temporal image data of VEGETATION sensor. *Remote Sensing of Environment*, 90, 477-489.

Borak, J.S. and Strahler, A.H. (1999), Feature selection and land cover classification of a MODIS-like data set for a semiarid environment. *International Journal of Remote Sensing*, 20, 919-938.

Boser, B.E., Guyon, I.M. and Vapnik, V.N. (1992), A training algorithm for optimal margin classifiers. In D. Haussler (Ed.), *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory* (pp. 144-152). ACM Press.

Bruzzone, L., Roli, F. and Serpico, S.B. (1995), An extension to multiclass cases of the Jeffries-Matusita Distance. *IEEE Transaction on Geoscience and Remote Sensing*, 33, 1318-1321.

Caetano, M., Carrão, H. and Painho, M. (2005), *Alterações da ocupação do solo em Portugal Continental: 1985 – 2000*. Amadora, Portugal: Instituto do Ambiente, 52 pp.

Carrão, H., Gonçalves, P. and Caetano, M. (2007), Use of intra-annual satellite imagery time-series for land cover characterization purpose. *EARSeL eProceedings*, 6, 1-11.

Cihlar, J. (2000), Land cover mapping of large areas from satellites: status and research priorities. *International Journal of Remote Sensing*, 21, 1093-1114.

Cristianini, N. and Shawe-Taylor, J. (2000), *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, 189 pp.

DeFries, R., Hansen, M., Townshend, J.R.G. and Sholberg, R. (1998), Global land cover classifications at 8km spatial resolution: The use of training data derived from Landsat imagery in decision tree classifiers. *International Journal of Remote Sensing*, 19, 3141-3168.

DeFries, R. and Belward, A.S. (2000), Global and regional land cover characterization from satellite data: an introduction to the Special Issue. *International Journal of Remote Sensing*, 21, 1083-1092.

Di Gregorio, A. and Jansen, L.J.M. (2000), *Land Cover Classification System*. Rome, Italy: FAO, 179 pp.

Duchemin, B., Guyon, D. and Lagouarde, J.P. (1999), Potential and limits of NOAA-AVHRR temporal composite data for phenology and water stress monitoring of temperate forest ecosystems. *International Journal of Remote Sensing*, 20, 895-917.

Foody, G. M. and Mathur, A. (2006), The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. *Remote Sensing of Environment*, 103, 179−189.

Foody, G. M., Mathur, A., Sanchez-Hernandez, C. and Boyd, D. S. (2006), Training set size require-ments for the classification of a specific class. *Remote Sensing of Environment*, 104, 1-14.

Gonçalves, P., Carrão, H., Pinheiro, A. and Caetano, M. (2006), Land cover classification with Support Vector Machine Applied to MODIS imagery. In A. Marçal (Ed.), *Global Developments in Environmental Earth Observation from Space* (pp. 517-526). Rotterdam: Millpress.

Guyon, I. and Elissee, A. (2003), An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3, 1157-1182.

Haertel, V. and Landgrebe, D. (1999), On the Classification of Classes with Nearly Equal Spectral Responses in Remote Sensing Hyperspectral Image Data. *IEEE Transactions on Geoscience and Remote Sensing*, 37, 2374-2386.

Hammond, T.O. and Verbyla, D.L. (1996), Optimistic bias in classification accuracy assessment. *International Journal of Remote Sensing*. 17, 1261-1266.

Hansen, M.C., DeFries, R.S., Townshend, J.R.G. and Sohlberg, R. (2000), Global land cover classification at 1km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, 21, 1331-1364.

Herold, M., Gardner, M.E. and Roberts, D.A. (2003), Spectral resolution requirement for mapping urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 1907–1919.

Ho, T.K. and Basu, M. (2002), Complexity Measures of Supervised Classification Problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24, 289-300.

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

Jackson, Q. and Landgrebe, D. (2001), An adaptive classifier design for high-dimensional data analysis with a limited training data set. *IEEE Transactions on Geoscience and Remote Sensing*, 39, 2664-2679.

Jensen, J.R. (1996), *Introductory Digital Image Processing: A Remote Sensing Perspective*, 2nd Ed. New Jersey: Prentice Hall Inc., 379 pp.

Jimenez, L. and Landgrebe, D. (1999), Hyperspectral Data Analysis and Feature Reduction Via Projection Pursuit. *IEEE Transactions on Geoscience and Remote Sensing*, 37, 2653-2667.

Johnson, R.A. and Wichern, D.W. (1998), *Applied Multivariate Statistical Analysis*, 4th ed. New Jersey: Prentice Hall, Upper Saddle River, 767 pp.

Knight, J.F., Lunetta, R.L., Ediriwickrema, J. and Khorram, S. (2006), Regional Scale Land-Cover Characterization using MODIS-NDVI 250 m Multi-Temporal Imagery: A Phenology Based Approach. *GIScience and Remote Sensing*, 43, 1-23.

Landgrebe, D. (2002), Hyperspectral Image Data Analysis as a High Dimensional Signal Processing Problem (Invited). *Special Issue of the IEEE Signal Processing Magazine*, 19, 17-28.

Landgrebe, D. (2005), Multispectral Land Sensing: Where From, Where To?. *IEEE Transactions on Geoscience and Remote Sensing*, 43, 414-421.

Lee, C. and Landgrebe, D. (1993), Analyzing High Dimensional Multispectral Data. *IEEE Transactions on Geoscience and Remote Sensing*, 31, 792-800.

Liu, J., Zhuang, D., Luo, D. and Xiao, X. (2003), Land-cover classification of China based on integrated analysis of AVHRR imagery and geo-spatial data. *International Journal of Remote Sensing*, 24, 2485 - 2500.

Loveland, T.R., Merchant, J.M., Ohlen, D.O. and Brown, J.F. (1991), Development of a land- cover characteristics database for the conterminous U.S.. *Photogrammetric Engineering and Remote Sensing*, 57, 1453-1463.

Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L. and Merchant, J.W. (2000), Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, 21, 1303–1330.

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

Matsuoka, M., Hayasaka, T., Fukushima, Y. and Honda, Y. (2004), Land Cover Classification Over Yellow River Basin Using Satellite Data. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 20-24 September, Fairbanks, USA.

Mather, P. M. (2004), *Computer Processing of Remotely-Sensed Images: An Introduction*, 3rd Ed. San Francisco, CA: John Wiley and Sons, Ltd., 324 pp.

Mercier, G. and Lennon, M. (2003), Support vector machines for hyperspectral image classification with spectral based kernels. In *Proceedings of IEEE International Geoscience & Remote Sensing Symposium (IGARSS 2003)*, 21-25 July, Toulouse, France.

Pal, M. and Mather, P.M. (2005), Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, 26, 1007-1011.

Painho, M., and Caetano, M. (2006), *Cartografia de Ocupação do Solo, Portugal Continental, 1985-2000*. Amadora, Portugal: Instituto do Ambiente, 56 pp.

Seeger, M. (2004), Gaussian processes for machine learning. *Interrnational Journal of Neural Systems*, 14, 1-38.

Shah, C.A., Watanachaturaporn, P., Arora, M.K. and Varshney, P.K. (2003), Some Recent Results on Hyperspectral Image Classification. In *IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, 27-28 October 2003, NASA Goddard Spaceflight center, Greenbelt, MD.

Singh, S. (2003), Multi-resolution estimates of classification complexity. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25, 1534-1539.

Scholkopf, B., Smola, A. J., Williamson, R. C. and Bartlett, P. L. (2000), New support vector algorithms. *Neural Computation*, 12, 1207–1245.

Stehman, S.V. and Czaplewski, R.L. (1998), Design and analysis for thematic map accuracy assessment: fundamental principles. *Remote Sensing of Environment*, 64, 331–344.

Suykens, J., Gestel, T., de Brabanter, J., de Moor, B. and Vandewalle, J. (2002), *Least Squares Support Vector Machines*. Singapore: World Scientific, 303 pp.

Thenkabail, P.S., Smith, R.B. and De Pauw, E. (2000), Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sensing of Environment*, 71, 158–182.

Thenkabail, P.S., Schull, M. and Turral, H. (2005), Ganges and Indus river basin land use/cover (LULC) and irrigated area mapping using continuous streams of MODIS data. *Remote Sensing of Environment*, 95, 317-341.

Van Dijk, A., Callis, S.L., Sakamoto, C.M. and Decker, W.L. (1987), Smoothing vegetation index profiles: an alternative method for reducing radiometric disturbances in NOAA/AVHRR data. *Phtogrammetric Engineering and Remote Sensing*, 53, 1059-1067.

Vapnik, V. (1982), *Estimation of dependencies based on empirical data*. New York: Springer Verlag.

Vapnik, V. (1998), *Statistical Learning Theory*. New York: John Wiley.

Vermote, E.F. and Vermeulen, A. (1999), Atmospheric correction algorithm: spectral reflectances (MOD09). Algorithm Technical Background Document, Available online at: http://modis.gsfc.nasa.gov/data/atbd/ (accessed 17 October 2005).

Vieira, C.A.O., Mather, P.M. and Aplin, P. (2001), Multitemporal classification of agricultural crops using the spectral-temporal response surface. In Lorenzo Bruzzone and Paul Smits (Ed.), *Analysis of Multi-Temporal Remote Sensing Images – Series in Remote Sensing, Vol. 2* (pp. 290-297). Singapore: World Scientific.

Watanachaturaporn, P. Arora, M.K. and Varshney, P.K. (2004), Evaluation of factors affecting support vector machines for hyperspectral classification. In *Proceedings of the American Society for Photogrammetry & Remote Sensing (ASPRS) 2004 Annual Conference*, 23-28 May 2004, Denver, CO.

Watanachaturaporn, P., Varshney, P.K. Arora, M.K. (2005), Multisource fusion for land cover classification using support vector machines. In *Proceedings of the Eighth*

International Conference on Information Fusion (FUSION 2005), 25-29 July 2005, Philadelphia, USA.

Wessels, K.J., De Fries, R.S., Dempewolf, J., Anderson, L.O., Hansen, A.J., Powel, S.L. and Moran, E.F. (2004), Mapping regional land cover with MODIS data for biological conservation:Examples from the Great Yellowstone Ecosystem, USA and Pará State, Brazil. *Remote Sensing of Environment*, 92, 67-83.

Wolfe, R.E., Nishihama, M., Fleig, A.J., Kuyper, J.A., Roy, D.P., Storey, J.C. and Patt, F.S. (2002), Achieving sub-pixel geolocation accuracy in support of MODIS land science. *Remote Sensing of Environment*, 83, 31-49.

Xiao, X., Braswell, B., Zhang, Q.Y., Boles, S., Frolking, S. and Moore, B. (2003), Sensitivity of vegetation indices to atmospheric aerosols: Continental-scale observations in Northern Asia, *Remote Sensing of Environment*, 84, 385-392.

Xiong X., Sun, J., Esposito, J., Guenther, B. and Barnes, W.L. (2003), MODIS Reflective Solar Bands Calibration Algorithm and On-orbit Performance. In *Proceedings of SPIE – Optical Remote Sensing of the Atmosphere and Clouds III*, 4891, 95-104.

Zhu, Z. and Evans, D.L. (1994), U.S. forest types and predicted percent forest cover from AVHRR data. *Phtogrammetric Engineering and Remote Sensing*, 60, 525-531.

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

4 SEPARABILITY ANALYSIS OF LAND COVER CLASSES AT REGIONAL SCALE: A COMPARATIVE STUDY OF MERIS AND MODIS DATA

Carrão, H., Sarmento, P., Araújo, A. and Caetano, M. (2007), Separability analysis of land cover classes at regional scale: a comparative study of MERIS and MODIS data. In *Proceedings of the ENVISAT Symposium* (unpaginated CD-ROM), 23-27 April 2007, Montreaux, Switzerland.

Abstract

Medium spatial resolution satellite images have been recurrently used for automatic land cover mapping of large geographical areas. Despite the innovative developments in temporal and spectral resolutions of these sensors, their spatial resolution still hampers the effective discrimination between detailed cover types within fragmented landscapes. In this study we exploit the ability of MERIS and MODIS imagery for the discrimination between land cover classes in Portugal. We aim at identifying the most suitable subset of classes, from a recently recommended land cover nomenclature, to be used in the regular production of medium spatial resolution land cover products at national scale. Final results show differences on the spectral separability of land cover classes with MERIS and MODIS images.

4.1. Introduction

The Remote Sensing Unit (RSU) of the Portuguese Geographic Institute (IGP) has recently recommended a land cover nomenclature to be used in the regular production of user driven land cover products for multi-scale environmental monitoring in Portugal. The adopted nomenclature, up to now entitled LANDEO [1], is composed by 25 comprehensive classes that were defined through the Land Cover Classification System (LCCS) from Food and Agriculture Organization (FAO) [2]. The rationale behind the development of this nomenclature was three-fold: (1) a nomenclature that is well adapted to the type of landscapes existent in regions with climatic characteristics similar to the Portuguese mainland; (2) a nomenclature that is compatible with established ones (e.g. CORINE Land cover, Global Land Cover and the International Geosphere-Biosphere Programme nomenclatures), in order to turn possible the comparison between our maps and others using different nomenclatures; and (3) a nomenclature that can be derived from satellite images with different spatial resolutions [1, 3].

In the framework of land cover mapping it is understandable that sensors collecting spectral data at high spatial resolutions, such as the Landsat Thematic Mapper (TM), cannot be used to derive comparable products at national and regional scales. This is due to incomplete spatial coverage, infrequent temporal coverage with inevitable cloud contamination and the associated large data volumes not practicable in an automatic land cover classification context [4]. Recently launched Earth Observation (EO) sensors, such as the MEdium Resolution Imaging Spectrometer (MERIS) and the Moderate Resolution Imaging Spectroradiometer (MODIS), exhibit enhanced spectral and temporal resolutions, wide geographical coverage and improved atmospheric correction. Thus, images collected by those sensors should help addressing the challenges of automatic land cover classification of broad geographical areas [3]. Moreover, [5] affirm that the availability of a large number of spectral bands makes it possible to identify more detailed land cover classes with higher accuracy than would be possible with data from elder sensors, such as the Advanced Very High Resolution Radiometer (AVHRR).

In this paper we evaluate the aptitude of medium spatial resolution imagery data, acquired by MERIS and MODIS sensors, for land cover characterization in Portugal. The Portuguese landscape is extremely fragmented and the use of medium spatial resolution satellite imagery for land cover cartography production is expected to be problematical in this context. At such a coarse resolution, land cover homogeneity is hardly guaranteed and many pixels are likely to contain features from numerous classes. Thus, the goal of this study was to identify the most suitable set of land cover classes from LANDEO nomenclature that can be accurately discriminated in Portugal through supervised classification approaches. We can separate the main objective in two specific goals: 1) evaluate the ability of high spectral resolution MERIS and MODIS images to discriminate properly between LANDEO classes; 2) recommend a combination of land cover classes to be merged in order to properly represent the Portuguese landscape at the spatial resolution of those satellite images.

4.2. Land cover nomenclature

The original LANDEO nomenclature is introduced in Table 4.1. The nomenclature comprises 25 land cover classes, but we just considered 17 classes in this study. Burnt areas and clear cuts were removed because they are dynamic classes that vary constantly along a one year period and cannot be attached to a specific moment in time. Moreover, we have already defined a set of procedures to accurately map burnt areas and clear cuts, namely those present in [6]. Water areas were also removed from the study. This cover type is easily distinguishable from other covers through automatic classification of MERIS and MODIS

images [3, 7]; as a consequence, its separability does not need to be evaluated. Mosaics represent complex areas that comprise features of several homogeneous classes in different proportions. The strategy for the classification of these areas involves *a posteriori* analysis of primary classification results of homogeneous classes. Thus, they were also discarded from this study, but are later considered in the discussion section.

Table 4.1. LANDEO land cover nomenclature, class label and respective number of collected sample observations.

Label	Land Cover Class	Sample size
11	Continuous Artificial Areas	44
12	Discontinuous Artificial Areas	42
21	Rainfed Herbaceous Crops	41
22	Irrigated Herbaceous Crops	43
23	Rice Crops	45
241	Permanent Evergreen Crops (Trees or Shrubs)	46
242	Permanent Deciduous Crops (Trees or Shrubs)	44
25	Mosaic Cultivated and managed lands	-
311	Broadleaved Closed Trees	49
312	Broadleaved Open Trees	44
321	Needleleaved Closed Trees	49
322	Needleleaved Open Trees	39
331	Mixed Closed Trees	43
332	Mixed Open Trees	38
34	Closed to Open Shrubland	45
35	Closed to Open Herbaceous Vegetation	41
36	Sparse Vegetation	39
37	Mosaic Trees/Shrubs/Herbaceous	-
38	Recently Burnt (Trees or Shrubs)	-
39	Clear-cuts	-
310	New plantations	-
41	Mosaic Cultivated and managed lands/ Natural and Semi-natural areas/Artificial areas	-
5	Permanent Wetlands	-
6	Barren	43
7	Water Bodies	-

4.3. Dataset

The data used to perform this study includes: 1) MERIS and MODIS satellite images for the year 2005; 2) orthorectified aerial images acquired during the years of 2004, 2005 and 2006, and covering the whole Portuguese territory – these images have the following characteristics: four spectral bands in the blue, green, red and infrared wavelengths, and 50cm spatial resolution; and 3) a set of reference sample observations to be used as auxiliary information in the definition of sufficient statistics, i.e. mean and sum-of-squares-and-cross-

products matrix, characterizing the reflectance of each land cover class within spectral wavelengths measured by both sensors. Reference sample is a set of geographic locations representing the true cover type at Earth's surface in those areas.

The description of MERIS and MODIS sensors and images features is important in order to properly understand the results attained in this study. Hence, next we describe the general characteristics of sensors and respective images, as well as the properties of the colleted reference sample.

4.3.1. Imagery data

MERIS sensor is on board of ENVISAT, an advanced polar-orbiting Earth Observation (EO) satellite launched by the European Space Agency (ESA) in 2002. This sensor allows a global coverage of the Earth in 3 days, and provides the most radiometrically accurate data on Earth's surface that is currently acquired from space [8]. MERIS data can be provided at 3 different levels of processing - Level 0, Level 1b, and Level 2 - and at 3 different spatial resolutions – Full (FR), Reduced (RR) and Low (LR). In this study we exploit Level 2 Full Resolution MERIS imagery. The selected product measures the solar radiation reflected by the Earth at a ground spatial resolution of 300m and consists of calibrated surface reflectances in 13 groups of wavelengths (original bands 11 and 15 were removed from Level 2 products since they address O_2 content and water vapour, respectively). These bands range from the visible to the near infrared wavelengths (390nm to 1040nm) and were atmospherically corrected for Rayleigh scattering, ozone, water vapour absorption and aerosol content [8]. These images contain additional geophysical parameters and quality flags provided on a pixel by pixel basis, which address the quality and the validity of the product.

MODIS is an instrument on board Terra and Aqua satellites from National Aeronautics and Space Administration (NASA). Terra MODIS and Aqua MODIS take between one and two days to cover the entire Earth's surface, with a complete 16-day repeat cycle. Both sensors acquire data in 36 spectral bands, or groups of wavelengths, and their spatial resolution (pixel size at nadir) is 250m for channels 1 and 2 (0.6µm - 0.9µm), 500m for channels 3 to 7 (0.4µm - 2.1µm) and 1000m for channels 8 to 36 (0.4µm - 14.4µm). These channels are calibrated on orbit by a solar diffuser (SD) and a solar diffuser stability monitor (SDSM) system, which convert the Earth surface radiance to radiometrically and geo-located calibrated products for each band [9].

Imagery data acquired by MODIS sensor is used to generate multiple products at different pre-process stages. In this study we exploit MOD09A1 imagery data. This specific product is

an estimate of the surface spectral reflectance imaged at a nominal spatial resolution of 500m for the first seven bands as it would have been measured at ground level if there were no atmospheric scattering or absorption. The applied correction scheme compensates for the effects of atmospheric gases, aerosols and thin cirrus clouds [10]. As similar as for the MERIS products, the MOD09A1 products have quality assurance data at the level of the individual pixel, at the level of each band and each resolution, and at the level of the whole file. MOD09A1 images were used in this study because they have a spatial resolution that is the most equivalent to that of MERIS Level 2 Full Resolution data. Still, it is important to refer that a pixel MODIS comprises an area of approximately three pixels MERIS.

The comparison of MERIS and MODIS data for land cover discrimination in Portugal involved the utilization of single date images from each sensor. The acquisition date corresponds to August 2005. This month was chosen because this is the period of the year that allows for a better separation among general land cover classes in this geographic region of transitional climate [7].

4.3.2. Reference sample

We focus our study on a reference sample comprising a set of i.i.d. observations collected for each land cover class of the LANDEO nomenclature (except for those removed classes). Sample observations, each corresponding to a MERIS pixel area, were selected for each land cover class using the CORINE Land Cover map for 2000 (CLC2000) [11] as strata. The selection of reference sample observations with this specific squared area was constrained by the specifications of an on-going project within RSU, and forced us to some additional pre-process of final sample, which is described in next chapter.

High spatial resolution EO data, namely the orthorectified aerial images acquired during the years of 2004, 2005 and 2006, were used as the base data source to recheck reference land cover labels of selected sample observations. Recheck of reference sample information was derived by the method described by [12], i.e. by visual interpretation of high resolution EO data overlaid with a co-registered 300m fishnet corresponding to the MERIS data grid. At such a coarse resolution, land cover homogeneity is hardly guaranteed for some projected classes and many pixels are likely to contain features from two or more distinct classes. Whenever it happened, we decided to assign these sampling units to the class with the area identified as dominant in the aerial image [13]. However, to avoid these occurrences, reference land cover observations were selected and identified mainly for geographical spots located in the centre of 3x3 MERIS pixels areas with homogeneous land cover. In Table 4.1 we presented the number of sample observations collected for each class.

Let us now introduce some notation that is required to comprehend the statistical analyses described in next section. For each sample observation we can associate a vector $\boldsymbol{x}$ of measurements on p spectral bands, where p=13 for MERIS and p=7 for MODIS. This vector represents a measurement of the spectral reflectances at Earth's surface within 300x300m and 500x500m areas imaged, respectively, by MERIS and MODIS sensors. Moreover, for each land cover class, we can define the sample mean multispectral vector $\bar{x}$ and the respective sum-of-squares-and-cross-products matrix $\boldsymbol{S}$.

4.4. Methodology

This section includes two main parts: 1) a description of the process used to improve the quality of the collected reference sample; and 2) an introduction to the statistical analysis performed to evaluate the separability of land cover classes on the basis of spectral information from the samples.

4.4.1. Sub-sample selection

The collected sample includes some observations that could not be used to properly evaluate the land cover classes' separability on the basis of reflectances measured by both sensors. Firstly, we remove from the final set all sample observations that were flagged "invalid" by MERIS and MODIS data, i.e. pixels that were identified has containing clouds, shadows and other inadequate features. Secondly, we remove all sample observations that did not correspond to an homogeneous land cover within a MODIS pixel area. The goal of this task was to guarantee that the collected observations were representative of each land cover class at the spatial resolutions of both MERIS and MODIS images. Because one pixel MODIS comprises almost three pixels MERIS and the Portuguese landscape is extremely fragmented, it was necessary to certify that sample observations collected for MERIS 300x300m pixels were representative of the same land cover within a MODIS 500x500m pixel area. Therefore, to reduce uncertainty in the reference labeling of observations at MODIS spatial resolution, we submitted the sample to a neighborhood analysis within a window of 3x3 MERIS pixels (note that this window is the minimum regular area that completely comprises a MODIS pixel if no geometric errors exist). Our approach consisted on detecting neighborhood pixels of sample observations, within one pixel radius, that do not seem to belong to the characteristic pattern of variability of the land cover class of the respective sample observation. One technique that can be used to detect "outlier" pixels in the neighborhood window is to compute the squared generalized multispectral distance from each neighbor pixel j to the mean multispectral vector $\bar{x}$ of the land cover class that characterizes the sample observation pixel:

$$d_j^2 = (x_j - \bar{x})^t S^{-1} (x_j - \bar{x}),\ j = 1,2,\ldots,8 \qquad (4.1)$$

where x_j is neighbor j of sample observation measured at MERIS p spectral bands. This technique is based on the assumption that $\bar{x}$ was generated from a multivariate normal distribution. The analysis of this assumption was performed previously for each land cover class using the method described by [14, pp. 199]. If this assumption is true, or approximately true, we can compare the results with the chi^2_p distribution at some 100(1- α) % confidence level. If d^2_j is higher than the $chi^2_p(\alpha)$, then we can reject the hypothesis that the neighbor pixel j belongs to the same land cover class of the sample pixel. We remove from the sample set all observations with less than six direct geographic neighbors that belong to the same land cover class at 95% confidence level. Nonetheless, this decision may still not be sufficient to avoid the existence of some noisy sample observations at the MODIS spatial resolution. Note that two MERIS pixels correspond to approximately 66% of a MODIS pixel area. In accordance, in the case of misregistration of MODIS images, there is a change that 66% of the sample observation area to be occupied by cover types different from that of interest. Accordingly, this difficulty may bias the spectral statistics of the respective land cover class, but this two pixel limit was necessary to guarantee that sample size per class was adequate for posterior statistical analyses.

4.4.2. Statistical separability analysis of land cover classes

The core of our study is to evaluate the ability of MERIS and MODIS imagery to discriminate between high detailed land cover classes that were theoretically defined to portray the main characteristics of the Portuguese landscape. One rational approach is to compare the sample mean reflectance vector $\bar{x}$ over each land cover class to see whether the spectral responses differ significantly over the classes [14]. Consider a random reference sample of size n_1 from land cover class ***1*** and a sample of size n_2 from land cover class ***2***. We want to make inferences about the difference between the mean reflectance vectors of class ***1*** and ***2***, i.e. μ_1 - μ_2, by taking into account sample information. For instance, we want to answer the question: is $\mu_1=\mu_2$? Or, equivalently, is μ_1 - $\mu_2=0$? The *Hotelling's* T^2 statistic can be used for testing H_0: μ_1 - $\mu_2=0$ and H_1: μ_1 - $\mu_2<>0$, and is given by:

$$T^2 = [\bar{x}_1 - \bar{x}_2]\left[\frac{1}{n_1}S_1 + \frac{1}{n_2}S_2\right]^{-1}[\bar{x}_1 - \bar{x}_2] \qquad (4.2)$$

where $\bar{x}_1$ and $\bar{x}_2$ are the sample mean reflectance vectors of classes ***1*** and ***2***, and S_1 and S_2 are the sum-of-squares-and-cross-products matrices of classes ***1*** and ***2***. T^2-statistic has a chi^2_p

distribution when $\bar{x}_1$ and $\bar{x}_2$ are approximately multivariate normal distributed [14]. At the α level of significance, we reject H_0 in favor of H_1 if the observed

$$T^2 > chi^2_p(\alpha) \qquad (4.3)$$

If H_0 is not rejected, we conclude that no average difference exists between two classes and they may not be properly separated at 100(1-α) % confidence level, i.e. from a Bayesian viewpoint.

Having this condition in mind, we developed an approach to group land cover classes for which no average difference exists at 95% confidence level. This approach was individually performed considering sample spectral information from MERIS (p=13) and from MODIS (p=7). Through an agglomerative hierarchical method that started with the individual 17 land cover classes of LANDEO nomenclature, we successively grouped two-by-two land cover classes when no average distance existed at 95% confidence level. At each time the most similar classes, i.e. those presenting the minimum T^2, were grouped until the computed statistic were greater then the respective chi^2_p distribution value for all pairs of classes. When all T^2 values were larger than the chi^2_p values, we can suggest that land cover groups differ significantly on the basis of the input spectral feature space. This escorts us to some proper land cover nomenclature to be used in the production of maps through automatic classification of medium spatial resolution of both satellite images in Portugal.

4.5. Results and discussion

A final set of 30 sample observations per class was guaranteed at the final process of sub-sample selection. This number is the minimum necessary to certify that hypothesis testing can be correctly derived and the obtained results are of confidence [14, pp. 199]. Moreover, using the method introduced in previous chapter for assessing normality, we assured that the mean vectors $\bar{x}$ characterizing each land cover class come from a multivariate normal distribution. This result also reinforces the following outcomes of this study.

Let us now proceed with the statistical separability analysis of LANDEO land cover classes on the basis of the spectral information derived from MERIS and MODIS images. In Table 4.2 we present the successive grouping of classes performed with spectral information from both sensors. Recall that successive grouping of land cover classes stopped when there was a confirmation in the average difference between the two most similar classes at 95% level of confidence.

Table 4.2. Grouped classes with no average difference at 95% level of confidence.

Group merge	MERIS	MODIS
1	"321,322"	"321,322"; "35,36"
2	"6,36"	"331,332"; "35-36, 6"
3	"6-36, 35"	"331-332, 321-322"; "35-36-6, 34"
4	-	"331-332-321-322, 35-36-6-34"

The results introduced in Table 4.2 reveal several similar properties between both sensors in discriminating land cover classes. We perceive that classes 321 and 322 are the closest in the *p*-dimensional feature space defined by both sensors images. This is due to the fact that 321 class (Needleleaved closed trees) has a low tree crown cover, which is near the percentage of occupation that defines the separation with 322 forests. Another expected merger was the one comprising the classes 6, 36 and 35. These classes are conceptually analogous, and they only differ on the cover percentage of herbaceous vegetation. Moreover, it is important to refer that these classes have similar geographical distributions, i.e. correspond to transitional areas in the Portuguese landscape and usually occur imbricate. Thus, at these spatial resolutions, it is unlikely that these classes can be individually mapped. This situation has already been reported by [3,7].

Let us now comment the differences on the successive merging of classes attained with both sensors' images. We were not able to make a distinction at 95% confidence level between 331 and 332 forests using MODIS data. Open mixed forests (332) are an uncommon cover type in Portugal and at MODIS spatial resolution the sample observations might be contaminated by surrounding areas of closed mixed stands. This is in agreement with the difficulty encountered in finding homogeneous areas of open Needleleaved and mixed forests to collect sample observations. Moreover, their linkage with 321 and 322 forests is also explicable. Mixed forests (331 and 332) can be located in small areas closed to Needleleaved forests. The consequence is evident: mixed areas of forest at MERIS spatial resolution are dominated by Needleleaved forests at MODIS spatial resolution. Even so, we suppose that the inferior spectral resolution of MODIS imagery can also be responsible for this mixing, not being able to accommodate the differences in the reflectance of pure and "contaminated" areas of needleleaved forests. In addition, note that extensive areas of broadleaved forests (311) in the territory correspond to continuous eucalyptus plantations that are cultivated for paper production. For these reason mixed forests are not confounded with broadleaved forests.

Another difference between the results attained with both sensors respects to the class 34 (Shrubland). MODIS imagery is not able to provide enough information to spread out the

observations of this class from the group "6-36-35". Two reasons exist for their high proximity within MODIS data: 1) width and centre location of MERIS spectral bands is more effective for their discrimination; and 2) once again the spatial resolution was determinant for class merging, because they are united geographically. The first supposition is in accordance with the results of [15], which has shown that narrow bands can outperform broad bands in quantifying biophysical vegetation characteristics. The second proposal is also acceptable. Due to the fragmentation of the Portuguese landscape the elements that constitute these classes are regularly comprised together within areas of 500x500m.

An unexpected result at the beginning of this study was the "fusion" of forests "331-332-321-322" with other natural areas ("35-36-6-34") within MODIS spectral feature space. According to our analyses, the reason might be due to the fact that these types of forests are extremely fragmented and the interstices are occupied by other natural features, such as "35-36-6-34". Hence, at MODIS spatial resolution, we expect elements of these classes to exist in the same pixel, thus making their separation difficult. A positive evaluation of LANDEO nomenclature is that it prevents this situation with mosaic class 37, which comprises all these detailed land cover classes when their separation is not possible at certain spatial resolutions.

Classes 311 and 312 do not have similar multivariate distributions within MERIS and MODIS spectral feature spaces because they represent completely different land cover types within the Portuguese territory. The first represents large plantations of eucalyptus species, used for paper production, while the second represents agro-forestry areas dominated by cork-tree species. Another important remark is that they do present high average distance to the group of natural classes, even at MODIS spectral resolution. The reason is that these forests do not have herbaceous and shrub vegetation under canopy; as such, the spectral response within these areas is completely different at MODIS spectral resolution.

Artificial and agricultural areas have specific spectral characteristics that distinguish them from all other land cover classes. Furthermore, they represent managed land cover types, occupying extensive areas, and thus facilitating their discrimination at MODIS spatial resolution.

4.6. Conclusion

The outcomes of this spectral separability study suggest that MERIS and MODIS imagery have similar adequacy for land cover characterization in Portugal, although MERIS images performed better. The classes that show less separability in both types of images are on one hand Barren (6), Sparse Vegetation (36) and Closed to Open Herbaceous Vegetation (35), and on the other hand Needleleaved Closed Trees (321) and Needleleaved Open Trees (322).

The lower separability between some LANDEO classes, observed with MODIS images, may be a consequence of the larger pixel size, which makes difficult the existence of pure pixels of some of these classes at this spatial resolution. This is particularly true between Needleleaved and Mixed Trees (“331-332, 321-322”), and between these and natural areas (“35-36-6-34”). Moreover, these results suggest that enhanced spectral resolution of MERIS allows a better discrimination between these LANDEO classes.

To overcome the observed spectral and spatial inability of MERIS and MODIS, we suggest that the application of LANDEO nomenclature to map fragmented landscapes, such as the Portuguese one, should be simplified according to the results of this study. Open Needleleaved and Mixed Trees may be removed from nomenclature because they represent small transitional areas between stable cover types, whereas not being representative of the Portuguese land cover. On the other hand, closed to open herbaceous vegetation classes are frequent within the Portuguese landscape and at the same time they have similar spectral reflectances. For these reasons we suggest to maintain these classes in the nomenclature, but aggregate them before or after image classification. Another strategy to reduce these problems is the use of mosaic classes already foreseen in LANDEO nomenclature. Additionally, there are quantitative approaches (e.g. based on vegetation indices) that can be used to map vegetation densities more efficiently at these spatial resolutions, instead of categorical classification techniques.

4.7. Acknowledgements

This study was carried out in the framework of the project "LANDEO – User driven land cover characterisation for multi-scale environmental monitoring using multi-sensor earth observation data (PDCTE/MGS/49969/2003)" funded by "Programa Dinamizador das Ciências e Tecnologias para o Espaço" from "Fundação para a Ciência e Tecnologia", and from the "Announcement of Opportunity for the Utilisation of ERS and ENVISAT Data" from European Space Agency (ESA). Research by Hugo Carrão was founded by the "Fundação para a Ciência e Tecnologia" (SFRH/BD/18447/2004).

4.8. References

[1] Araújo, A. & M. Caetano (2006). Nomenclatura de Ocupação de Solo LANDEO, Portuguese Geographic Institute (IGP), Lisbon, Portugal.

[2] Di Gregorio, A. & Jansen, L.J.M. (2000). Land Cover Classification System, Rome, Italy: FAO, 179 pp.

[3] Carrão, H., Gonçalves, P. & M. Caetano (2006). MERIS Based Land Cover Characterization: A Comparative Study. In Proceedings of the ASPRS 2006 Annual Conference - Prospecting for Geospatial Information Integration, Reno, USA.

[4] Defries, R. & A. Belward (2000). Global and regional land cover characterization from satellite data: an introduction to the Special Issue. *International Journal of Remote Sensing*, 21 (6 & 7), 1083–1092.

[5] Jackson, Q. & D. Landgrebe (2001). An adaptive classifier design for high-dimensional data analysis with a limited training data set. *IEEE Transactions on Geoscience and Remote Sensing*, 39, 2664-2679.

[6] Armas, R. & M. Caetano (2005). Mapping changes in forest cover using multi-scale MODIS imagery. In Proceedings of the 31st International Symposium on Remote Sensing of Environment, 20 - 24 June 2005, Saint Petersburg, Russian Federation, unpaginated CD-ROM.

[7] Gonçalves, P., Carrão, H., Pinheiro, A. & Caetano, M. (2006). Land cover classification with Support Vector Machine Applied to MODIS imagery. In A. Marçal (Ed.), Global Developments in Environmental Earth Observation from Space (pp. 517-526). Rotterdam: Millpress.

[8] Curran, P.J. & Steele, C.M. (2005). MERIS: the re-branding of an ocean sensor. International Journal of Remote Sensing, 26(9), 1781-1798.

[9] Xiong X., Sun, J., Esposito, J., Guenther, B. and Barnes, W.L. (2003), MODIS Reflective Solar Bands Calibration Algorithm and On-orbit Performance. In Proceedings of SPIE – Optical Remote Sensing of the Atmosphere and Clouds III, 4891, 95-104.

[10] Vermote, E.F. & Vermeulen, A. (1999). Atmospheric correction algorithm: spectral reflectances (MOD09). Algorithm Technical Background Document, Available online at: http://modis.gsfc.nasa.gov/data/atbd/ (accessed 17 October 2005).

[11] Painho, M. & M. Caetano, Cartografia de Ocupação do Solo, Portugal Continental, 1985-2000. Amadora, Portugal: Instituto do Ambiente, 2006.

[12] DeFries, R., Hansen, M., Townshend, J.R.G. & Sholberg, R. (1998). Global land cover classifications at 8km spatial resolution: The use of training data derived

from Landsat imagery in decision tree classifiers. *International Journal of Remote Sensing*, 19, 3141-3168.

[13] Hansen, M.C., DeFries, R.S., Townshend, J.R.G. & Sohlberg, R. (2000). Global land cover classification at 1km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, 21, 1331-1364.

[14] Johnson, R.A. & Wichern, D.W. (1998). *Applied Multivariate Statistical Analysis, 4th ed.*, New Jersey: Prentice Hall, Upper Saddle River, 767 pp.

[15] Thenkabail, P.S., Smith, R.B. & De Pauw, E. (2000). Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sensing of Environment*, 71, 158–182.

5 MERIS BASED LAND COVER CHARACTERIZATION: A COMPARATIVE STUDY

Carrão, H., Gonçalves, P. and Caetano, M. (2006), MERIS Based Land Cover Characterization: A Comparative Study. In *Proceedings of the ASPRS 2006 Annual Conference - Prospecting for Geospatial Information Integration* (unpaginated CD-ROM), 1-6 May 2006, Reno, Nevada, USA.

Abstract

Knowledge of land cover spatial distribution is important for many activities, including environmental monitoring, land planning, and resource management. Land cover information at the appropriate time and over large geographical regions can only be derived from satellite images. With the recent launch of MERIS, a wide range of new possibilities for the periodic land cover characterization at regional scale is available. This sensor offers a combination of innovative features, such as high spectral and temporal resolutions, wide geographical coverage and improved atmospheric correction. We believe that the exploitation of data obtained by this new sensor fills previous technological gaps, improving automatic land cover classes' discrimination. At the same time, the extra spectral information provided by MERIS can introduce some difficulties on land cover characterization, since the dimensionality of feature space can be overloaded. In this paper we report the comparison of Support Vector Machine (SVM), a new supervised classification system that is independent of feature space dimensionality, with most frequently used supervised classifiers for the land cover classification of high dimensional MERIS imagery. To reduce feature space dimensionality and filter out redundant information a Fisher's Discriminant Analysis (FDA) was performed for spectral data rearrangement. The original and rearranged/reduced data sets were then used as inputs for the several classifiers. The results of the classification allow identifying the efficiency of SVM and FDA to improve land cover classification of high dimensional imagery data. The study was carried out with MERIS Full Resolution data from 2004 for the continental Portuguese territory.

5.1. Introduction

Land cover cartography production is an essential task for scientific research and earth science applications, such as landscape management and biodiversity assessment, as well as,

in general, to support environmental, social, and economic policies. Therefore, mapping landscape regimes that are persistently changing under the influence of several factors (e.g., forest fires, clear cuts, and urban fabric growth) must be a regular process. Land cover maps become rapidly out of date, while the need for updated intermediate-scale cartography is constantly increasing to help near real-time spatial decision-making. Remote sensing data are attractive for deriving land cover cartography by means of digital image classification, because updated information for continuous time periods and over extensive geographical regions can only be derived from satellite systems.

Visual interpretation and manual classification of remote sensing data is reliable and accurate, but cannot assure any longer the production in time of land cover knowledge that is required within many applications, given that is time consuming and economically expensive. To cope with this situation, several supervised and unsupervised algorithms have been developed and successively implemented for land cover classification of multi-spectral satellite imagery (Shah *et al.*, 2003). Even though the automatic exploitation of satellite data for land cover cartography production has proved to be cost-effective, some difficulties associated with imagery characteristics and classification specificities still hamper their effective employment (Gonçalves *et al*., 2005): (1) different land covers share similar spectral signatures and cannot always be automatically separated using classical remote sensing data; (2) conventional image classification techniques fail at separating a large number of land cover classes with acceptable misclassification rates.

Recently launched Earth Observation (EO) sensors, such as the MEdium Resolution Imaging Spectrometer (MERIS) and the Moderate Resolution Imaging Spectroradiometer (MODIS), exhibit enhanced spectral and temporal resolutions and offer new potentials and challenges to data analysis. Jackson and Landgrebe (2001) affirm that the availability of a large number of spectral bands makes it possible to identify more detailed land cover classes with higher accuracy than would be possible with the data from earlier sensors. Quite as well, several studies have proved the advantages of performing a land cover classification based on multi-temporal satellite imagery data (e.g., Gonçalves *et al*., 2005; Price *et al*., 2002; Maxwell *et al*., 2002; Vieira *et al*., 2001). However, the exploitation of spectral and temporal characteristics of MERIS imagery data, for the improvement of land cover spatial and thematic characterization, is still of confidential use in remote sensing community. Presently, few research studies and operational programs are focused on the analysis of MERIS data for land cover cartography production (e.g., Clevers *et al*., 2004; Gessner *et al*., 2004; Santos *et al*., 2005).

The introduction of these new generation sensors resulted in a large increase of remote sensing data volumes available for land cover classification purposes. Still, high dimensional remote sensing imagery provides a challenge to the current classification techniques. Existing algorithms often fail to deliver high classification accuracies and tend to suffer from the problem of the "curse of dimensionality", many times referred to as the Hughes phenomenon (Hughes, 1968). Although it is intuitive to think that increasing the dimension of the features should never reduce the classification performance, since we are providing a larger amount of information, the performance can in fact decrease while we feed more data to the system (Guyon and Elissee, 2003). A large number of land cover classes of interest and a large number of spectral bands available require numerous training samples for classification task that not always are available, because are expensive and tedious to acquire (Jackson and Landgrebe, 2001). Thus, the development of new methods to analyze high dimensional satellite imagery data became extremely necessary (Haertel and Landgrebe, 1999; Landgrebe, 2002; Shah *et al.*, 2003). As a result, the dimension reduction of satellite imagery data, via a pre-processing transformation technique that maps data from a high order dimension to a low order dimension, is required. The goal of dimension reduction, by feature extraction or selection, is to obtain a few features that discriminate the necessary classes with a high degree of accuracy (Guyon and Elissee, 2003). Principal Component Analysis (PCA) and Fisher's Discriminant Analysis (FDA) are dimension reduction techniques that could substantially increase classifiers performance by filtering out redundant information. This pre-classification procedure must improve classification accuracy of high dimensional imagery by means of frequently used classifiers.

A new supervised classification system based on the statistical learning theory (Vapnik, 1998), defined as Support Vector Machine (SVM), has recently been applied to the problem of high dimensional remote sensing data classification (Marçal *et al.*, 2005; Mercier and Lennon, 2003; Huang *et al.*, 2002; Zhu and Blumberg, 2002). This technique is said to be independent of the dimensionality of feature space as the main idea behind this classification technique is to separate the classes with a surface that maximizes the margin between them (Pal and Mather, 2005). SVM can obtain a better classification performance and have a more generalization capacity than classifiers that aim to minimize the training error rate alone, such as neural network classifiers (Shah *et al.*, 2003). Besides, a small number of training samples is enough to train the algorithm, which reveals interesting properties for high dimensional image processing (Mercier and Lennon, 2003).

In this paper we exploit several supervised pixel based techniques for the land cover classification of MERIS high dimensional imagery data. This is a preliminary study in the

framework of an on going research work that aims at developing a systematic classification methodology for the regular production of land cover cartography from medium spatial resolution satellite imagery. Specifically, the main goal of this study is to assess the utility of SVM for the land cover classification of MERIS high dimensional imagery data in comparison with most frequently used k-NN, K-means and LDC classifiers. Since these current classifiers are not well adapted to the classification of high dimensional data, we perform a dimension reduction of the original spectral information with Fisher's Discriminant Analysis and in addition appraise the accuracy of the several classification methods with the compressed information. To completely evaluate the ability of the high spectral resolution MERIS data for land cover characterization, we seek for the separation between 19 land cover classes by means of the several classifiers. We experimentally evaluate the performance of each classifier on a MERIS image of Portugal Continental for the date of August 14th, 2004.

5.2. Study area and data set

The study area is the entire Portuguese mainland territory. Portugal is in a transition zone featuring diverse landscapes representing both Mediterranean and Atlantic climate environments. This landscape heterogeneity allows for the extrapolation of the developed methodologies and of the obtained results to many other regions of the world.

5.2.1. MERIS data

The satellite imagery used in this study was acquired by the MEdium Resolution Imaging Spectrometer (MERIS). This sensor is on board of ENVISAT, an advanced polar-orbiting Earth Observation (EO) satellite launched by the European Space Agency (ESA) in 2002. MERIS is a 68.5° field-of-view push-broom imaging spectrometer that covers a swath width of 1,150km and allows a global coverage of the Earth in 3 days. This sensor measures the solar radiation reflected by the Earth at a maximum ground spatial resolution of 300m, in 15 spectral bands or groups of wavelengths, ranging from the visible to the near infrared (390 nm to 1040 nm) radiance, and provides the most radiometrically accurate data on Earth surface that is currently acquired from space (Curran and Steele, 2005). One of the most outstanding features of MERIS is the programmability of its spectral bands in their width and position.

MERIS data can be provided at 3 different levels of processing - Level 0, Level 1b, and Level 2 - and at 3 different spatial resolutions – Full (FR), Reduced (RR) and Low (LR). FR products have a spatial resolution of 300m, RR of 1200m and LR of 4800m. The Level 0 product corresponds to the data in its most raw format and will not generally be available to

the end users. Level 1b is annotated engineering data that has been calibrated radiometrically and geolocated – this is the Top-Of-the-Atmosphere (TOA) calibrated radiance measured in *mWm-2sr-1nm-1*. Level 2 products contain both geolocated geophysical parameters and surface radiance/reflectance depending on the surface type. In addition to the core geophysical data, MERIS Level 2 products contain geometric information, data describing the sun and viewing geometry, terrain height, some meteorological data and several flags that address the quality and the validity of the image.

In this study we exploit the Level 2 Full Resolution MERIS imagery. These data consists of calibrated surface reflectances in 13 groups of wavelengths (original bands 11 and 15 were removed from this product since they address O_2 content and water vapor) that are atmospherically corrected for Rayleigh scattering, ozone, water vapor absorption and aerosol content. We make use of a single image that was acquired by 14 August of 2004. This date was chosen due to several criteria. First, and according to Gonçalves *et al*. (2005), this is the period of the year that allows for a better separation among general land cover classes in this region of transitional climate; secondly, this image is the one that presents the minor cloud coverage among a set of 12 MERIS images acquired for Portugal for the year of 2004.

The MERIS original data was exported from raw format and converted to GeoTIFF using the BEAM VISAT 3.4 ® software application. Also, in order to combine these data with already existing auxiliary information it was necessary to project the image into a proper map projection (Hayford-Gauss, Datum Lisboa). This was easily accomplished since MERIS Level 2 imagery are already geolocated.

5.2.2. Ancillary data

To help the selection of samples for algorithm training and testing we made use of existent land cover databases and high spatial resolution earth observation data, namely:

- CORINE Land Cover map for 2000 (CLC2000) (Painho and Caetano, 2005) – land cover map with 44 classes derived from visual interpretation of Landsat-7 ETM+ images acquired in the summer of 2000;
- National cover of SPOT High Resolution Geometric (HRG) images acquired in 2003, owned by the Portuguese Geographic Institute (IGP);
- National cover of orthorectified colour infrared aerial photography for 1995;
- National Forest Inventory for 1995.

5.3. Methodology

5.3.1. Land cover nomenclature

In this study we used the land cover nomenclature developed by the Remote Sensing Unit of the Geographic Portuguese Institute (IGP) (Table 5.1). The nomenclature classes were defined through the Land Cover Classification System (LCCS) from Food and Agriculture Organization (FAO) (Di Gregorio, 2000). The description of land cover classes is based on LCCS classes' definition. The rationale behind the development of the nomenclature was two-fold: (1) a nomenclature that is well adapted to the type of landscapes existent in regions with characteristics similar to the Portuguese mainland, and (2) a nomenclature that is compatible with established ones (e.g., CORINE Land cover, Global Land cover and the International Geosphere-Biosphere Programme nomenclatures) in order to make possible the comparison of our maps with the ones that used other nomenclatures.

5.3.2. Land cover sampling

In order to test the several classification approaches using the defined nomenclature, we carry out a collection of samples for the 19 nomenclature classes for the year of 2004. The sampling process was performed by visual interpretation of high spatial resolution Earth Observation data, namely SPOT5-XS images acquired during the year of 2003 and orthorectified colour infrared aerial photography of 1995. To help sampling collection we also made use of the land cover databases described in the Study Area and Dataset section.

Each sample represents a specific land cover class covering at least 90% of the area covered by a MERIS 300-by-300 m pixel (same as the nominal resolution of used satellite images). The goal was to have, for each land cover class, the samples distributed all over the mainland territory in order to obtain a high-quality representation of possible regional within class differences. For each class we collected 40 sample observations and their spatial distribution is shown in Figure 5.1.

5.3.3. Dimension reduction

When data to be classified lie in multi-dimensional vector spaces, it is often the case that not all dimensions show the same discrepant power. Ever because components are highly correlated, or simply because all classes merge in a unique cluster, when looking at marginalized data in some particular direction, redundant or information less coordinates can be discarded from a classification viewpoint. As computational cost of classifiers increases

Table 5.1. Land Cover nomenclature and classes description.

Land Cover Class	Acronym	Description
Continuous Artificial Areas	11	The land cover consists of non-linear built up areas which can be further specified into industrial area(s) or urban area(s). The density of the impermeable surface(s) is high.
Discontinuous Artificial Areas	12	The land cover consists of non-linear built up areas which can be further specified into industrial area(s) or urban area(s). The density of the impermeable surface(s) is medium or low.
Rain fed Herbaceous Crops	21	Field(s) are covered by rain fed herbaceous crops. The crop covers the land during the cultivation period of a shifting system
Irrigated Herbaceous Crops	22	Field(s) are covered by irrigated herbaceous crops. The irrigation systems commonly used are surface, sprinkler and drip irrigation. The crop covers the land during the cropping period of a fallow system.
Rice Crops	23	Field(s) are covered by herbaceous crops. The crop uses infiltrated water as reserve for establishment. The crop covers the land during the cropping of a fallow system.
Permanent Evergreen Crops (Trees or Shrubs)	241	Evergreen tree crops cover a defined area. // Evergreen shrub crops cover a defined area.
Permanent Deciduous Crops (Trees or Shrubs)	242	Deciduous tree crops cover a defined area. // Deciduous shrub crops cover a defined area.
Broadleaved Closed Trees	311	The main layer consists of broadleaved closed trees. The crown cover is more than (70-60)%. The height is in the range 30-3 m.
Broadleaved Open Trees	312	The main layer consists of broadleaved woodland. The crown cover is between (70-60) and (20-10)%. The height is in the range 30-3 m.
Needleleaved Closed Trees	321	The main layer consists of needleleaved closed trees. The crown cover is more than (70-60)%. The height is in the range 30-3 m.
Needleleaved Open Trees	322	The main layer consists of needleleaved woodland. The crown cover is between (70-60) and (20-10)%. The height is in the range 30-3 m.
Mixed Closed Trees	331	The main layer consists of broadleaved closed trees. The crown cover is more than (70-60)%. The height is in the range 30-3 m. / The main layer consists of needleleaved closed trees. The crown cover is more than (70-60)%. The height is in the range of 30-3 m.
Shrubland	35	The main layer consists of closed to open shrubland. The crown cover is between 100 and 15% The height is in the range of 5-0.3 m.
Natural grassland	37	The main layer consists of closed to open herbaceous vegetation. The crown cover is between 100 and 15%. The height is in the range of 5-0.3 m.
Sparse Vegetation	38	The main layer consists of sparse herbaceous vegetation. The crown cover is between (20-10) and 1%. // The main layer consists of sparse shrub vegetation. The crown cover is between (20-10) and 4%. // The main layer consists of sparse tree vegetation. The crown cover is between (20-10) and 4%.
Recently Burnt (Trees or Shrubs)	310	The main layer consists of closed to open trees affected by forest fires. The crown cover is between 100 and 15%. The height is in the range of 30-3m. // The main layer consists of closed to open shrubland affected by forest fires. The crown cover is between 100 and 15%. The height is in the range of 5 - 0.3m.
Permanent Wetlands	5	Primarily vegetated areas containing more than 4 percent vegetation during at least two months a year. The environment is influenced by the presence of water over extensive periods of time, i.e. water is present for more than three months a year and when water is present less than three months a year, it is present 75 percent of the flooding time. The vegetative cover is characterized by the presence of (semi)natural vegetation which species composition, its environmental and ecological processes are indistinguishable from, or in a process of achieving, its undisturbed state. The vegetative cover is not artificial and does not need to be managed nor maintained. This class includes floating vegetation but excludes areas with only occasional flooding.
Barren	6	Primarily non-vegetated areas containing less than four percent vegetation during at least 10 months a year. The environment is influenced by the edaphic substratum. The cover is natural. Included are areas like bare rock and sands.
Water Bodies	7	The land cover consists of natural water bodies. A further specification can be made in flowing or standing water. // The land cover consists of artificial water bodies. A further specification can be made in flowing or standing water.

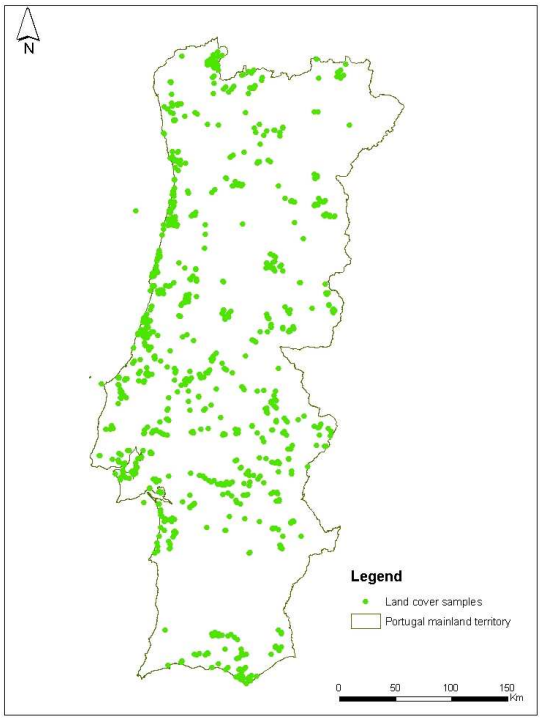

Figure 5.1. Spatial distribution of collected samples.

rapidly with the feature space dimension, it is in general profitable to keep the dimensionality of the problem as small as possible, as long as it does not degrade the classification accuracy. The dimensionality issue can turn even worse in the supervised case: if the feature space dimension is excessive compared to the size of the available training set, the classifier can fail at finding a simple decision rule. All this makes data dimension reduction an important step in classification task.

Feature selection and feature extraction are the two common ways to perform dimensionality reduction (Guyon and Elissee, 2003). Feature selection refers to techniques that identify the best subspace within the input features space, whereas extraction methods aim at finding the best transformation (by means of coordinate's combinations) of the input feature space. The choice between those two approaches depends on the physical meaning of the coordinate's dimension.

Principal Component Analysis (PCA) (Thomaz *et al.*, 1999; Jain *et al.*, 2000) is a widely used feature extraction that relies on the eigenvectors decomposition of the data covariance matrix. The corresponding linear transformation is an axis rotation that yields a system of systematically uncorrelated vectors (and statistically independent if data are normally distributed), aligned with the directions of maximum overall data variance. Mapped in this

new coordinate system, data are compactly represented and it is likely that retaining only the few first components associated to the largest eigenvalues, suffices to explain most of the overall data variability. Clearly though, this is an unsupervised method that does not take into account any prior information about classes, and therefore nothing guarantees that the directions of maximum variance are also the directions of maximum classes discrepancy.

Fisher's Discriminant Analysis (FDA) is a supervised extension of PCA, which identifies a reduced dimension system that enhance inter-classes separability (Mousavi *et al*., 2003; Jain *et al.*, 2000). Inter-classes separation is here measured by the Fisher criterion, calculated as the eigenvalues of the between-class scatter matrix (*B*) divided by the within-class scatter matrix (*W*). The within-class scatter matrix is defined by:

$$W = \sum_{j=1}^{c} \sum_{\mathbf{x_i} \in E_j} (\mathbf{x_i} - \overline{\mathbf{x_j}})^{t} (\mathbf{x_i} - \overline{\mathbf{x_j}}) \qquad (5.1)$$

where *x* denotes the vector data, *c* is the number of classes, and $\overline{\mathbf{x_j}}$ is the mean vector of class *j*. The between-class scatter is defined by:

$$B = \sum_{j=1}^{c} q_j (\overline{\mathbf{x_j}} - \overline{\mathbf{x}})^{t} (\overline{\mathbf{x_j}} - \overline{\mathbf{x}}) \qquad (5.2)$$

where q_j is the number of sample observations in class *j*, and $\overline{\mathbf{x}}$ is the mean of the overall data distribution.

The representation maximizing classes discrepancy was proposed by Fisher as the projection onto the eigenvectors associated to the largest eigenvalues of the matrix $W^{-1}B$.

5.3.4. Classification

k-Nearest Neighbor

k-NN is an instance-based learning algorithm (Cover and Hart, 1967) based on a distance function (e.g. Euclidian distance) between pairs of observations. The algorithm is quite simple: given a feature test vector, the system finds the k-nearest neighbors among the training vectors, and uses the labels of these k neighbors to determine the class of the unknown sample.

K-means

The decision rule adopted in K-means classifiers is to attribute a sample to-be-classified to the class which centre of gravity is closest.

Linear Discriminant Classifier

This classifier combines a Fisher Discriminant Analysis with a linear (or quadratic) separator function. Moreover, LDC implicitly assumes that data are normally distributed within each class (Johnson and Wichern, 1998).

Support Vector Machines

Support Vector Machines (SVM) are a new generation of supervised learning systems based on recent advances in statistical learning theory (Cristianini and Shawe-Taylor, 2000). Pioneered by the work on learning strategy by Vapnik and collaborators (Boser *et al.*, 1992; Vapnik, 1998), they have rapidly and successfully been applied to numerous real-world classification problems.

Conceptually, SVM rationale is that a classification problem that does not have a satisfying solution in its own observation space may have one simple and efficient in a more complicated representative system. As so, SVM use a hypothesis space of linear indicator functions to draw classification in a high dimensional (possibly infinite) feature space, image of the observation space by a non-linear mapping Φ.

Consider a binary classification-learning task[1] and the following data training set:

$$S\{(\mathbf{x}_1, y_1),...,(\mathbf{x}_N, y_N)\}, \mathbf{x}_i \in \mathbf{X} \subseteq \Re^{n_x}, y_i \in \{-1,+1\} \tag{5.3}$$

with samples $(\mathbf{x}_i, y_i)$ drawn i.i.d. according to some unknown but fixed probability distribution $P(\mathbf{x}, y)$. The non-linear mapping Φ transforms the n_x-dimensional input space $\mathbf{X}$ into the feature space $\mathbf{F}$, and let the set of hypotheses be linear hypothesis functions of the type:

$$h(\mathbf{x}) = \sum_{i=1}^{N} \alpha_i y_i \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}) \rangle + b \tag{5.4}$$

One remarkable fact about this hypothesis function (written here in its dual form) is that it implies the data only through their inner products in the feature space. Therefore, if Φ is properly chosen so that it obeys the so-called *kernel* condition,

$$\langle \Phi(\mathbf{x}), \Phi(\mathbf{z}) \rangle = K(\mathbf{x}, \mathbf{z}), \tag{5.5}$$

[1] For the sake of simplicity we restrict this introductory study to the case of a two-classes classification. Generalization to multiple classes problems is straightforward using a one-versus-the-rest strategy.

we do not even need to know the underlying feature map Φ to be able to learn in the feature space.

Assuming so, the problem is then to determine the coefficients $\{\alpha_i, i=1,...,N\}$ that minimize the classification error on unseen samples. Ideally, the best classifier should minimize the expected value of the *loss* or *risk*:

$$R(h) = \mathrm{E}_{P(\mathbf{x},y)}\{L(y,h(\mathbf{x}))\} = \int L(y,h(\mathbf{x}))dP(\mathbf{x},y), \qquad (5.6)$$

where the loss function $L(y,h(\mathbf{x})) \geq 0$, penalizes the deviations. In practice however, the joint probability function $P(\mathbf{x},y)$ is unknown, and the expected loss (also called *generalization loss*) is approximated by the *empirical classification error* based on the available information (i.e. the training set). Then, the empirical risk functional reads:

$$R_{emp}(h) = \sum_{i=1}^{N} L(y_i, h(\mathbf{x}_i)) . \qquad (5.7)$$

Often though, finding the hypothesis h^* that minimizes this empirical risk leads to an ill-posed problem that results in the well-known phenomenon described as *over-fitting* in the literature (i.e. the selected hypothesis is too complex). One way to avoid over-fitting tolerating noise and outliers, is to restrict the complexity of the hypothesis function, by introducing a regularization term (Vapnik, 1982). This regularization term is closely related to the notion of margin, another important concept for SVM that reflects the sensitivity and tolerance of the classifier to the samples $\{\mathbf{x}_i, i=1,...,N\}$ that stand "close" to the (non-linear) separator. Those points are called support vectors, and the solution that minimizes the regularized empirical risk function is referred to as a soft margin SVM (in contrast to a hard margin SVM that systematically zeroes the empirical loss). Solution to this quadratic optimization problem is classically obtained by Lagrangian theory and comes out to find the Lagrangian multipliers α^* that maximize the following quantity:

$$W(\alpha) = \sum_{i=0}^{N} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{N} y_i y_j \alpha_i \alpha_j (K(\mathbf{x}_i, \mathbf{x}_j) + \frac{1}{C}\delta_{ij}) , \qquad (5.8)$$

subject to $\sum_{i=1}^{N} y_i \alpha_i = 0$ and $\alpha_i \geq 0$, i=1,…N.

C is the regularization parameter that bounds the Lagrangian multipliers (i.e. the weights associated to the support vectors) controlling this way the capacity of the hypothesis function class[2].

Finally, a Radially Basis Function (RBF) kernel, $K(\mathbf{x},\mathbf{z}) = e^{-\frac{\|\mathbf{x}-\mathbf{z}\|^2}{\sigma^2}}$, with user pre-defined sample variance σ^2, is chosen because it maps the input space into an infinite dimensional feature space and often yields good results for nonlinear regression (Suykens *et al.*, 2002; Seeger, 2004).

5.4. Results and discussion

All classification results we report in this section were obtained using a cross validation technique that permits evaluating classifiers' efficiency (in terms of generalized bias) using only the collected sample set. We fixed to five the number of cross validation folds. Thus, the initial set of 40 samples per class was divided randomly into five subsets: each time four subsets are used for training purpose, while the remaining fifth is used for classification test.

Figure 5.2 presents the mean spectral signatures of land cover classes derived from the MERIS Level 2 image of August 14th, 2004. All curves, except those for impervious land covers like urban or barren, display a minimum and very similar reflectance value along the visible wavelength range that is comprised between 400 and 700nm. This is mainly because these reflectance wavelengths are scattered by the atmosphere and at the same time largely absorbed by chlorophyll – land cover pixels with vegetation do not show up very brightly in this fraction of the spectrum. Although spectral similarities exist among vegetated land cover classes, these bands can be useful for soil/vegetation discrimination and mapping of man-made features. On the other hand, near infra-red wavelengths are good for mapping vegetation biomass content, providing an evident contrast between different types of vegetation. Vegetation classes show a steep slope between the red and the near infra-red reflectance that is characterized by a peak around 700nm. This portion of the spectrum appears as disjunction point amongst vegetated land cover classes that maximizes their separation. This suggests that the combination and exploitation of land cover spectral profiles along visible and near-infrared wavelengths is important for the separation of

---

[2] There exists an alternative way for controlling the hypothesis function capacity, which amounts to fix the proportion of training samples that will lie between the boundary and the margin hyperplane. In SVM literature, this regularization strategy is often referred to as the ν-parameterization, and we will use this in our study.

different land cover classes and should significantly improve classification scores obtained with a lower dimensional representation of the available spectral data.

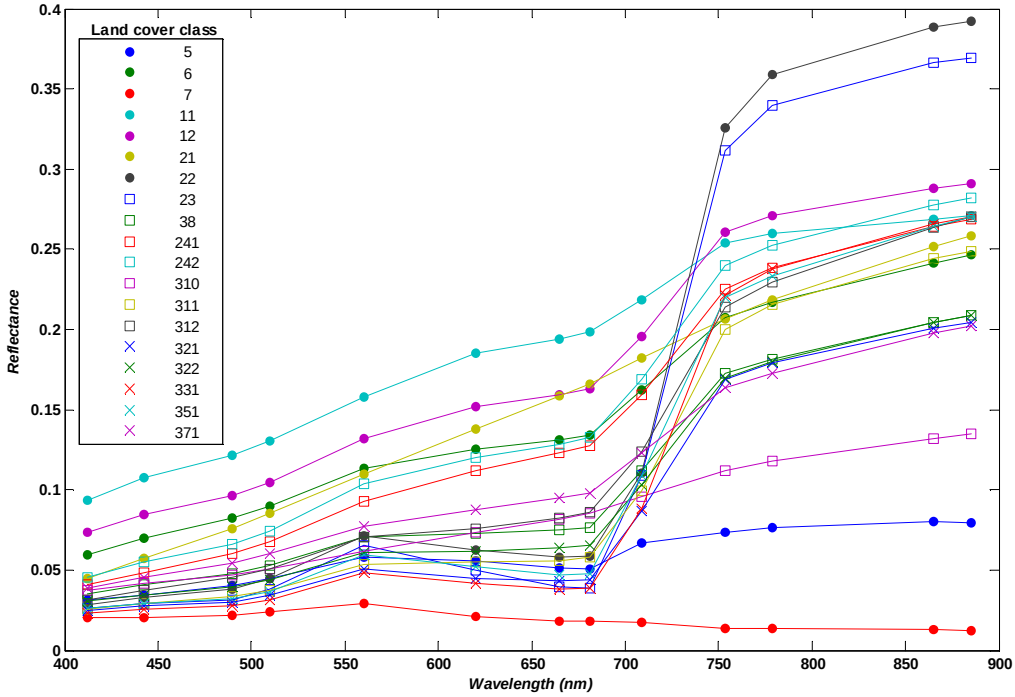

Figure 5.2. Mean spectral signatures for land cover classes derived from the MERIS image of August 14th, 2004.

We first performed an empirical estimation of the MERIS spectral bands correlations, which unquestionably evidenced the need of dimension reduction to remove redundant information. The averaged correlation coefficient between bands of the visible interval is about 95%, and it goes up to 99% for infra-red bands. In contrast, this same averaged coefficient falls down to 33% when measuring the correlation between visible and infra-red wavelengths. Then, it is not surprising that a Principal Component Analysis applied to the samples set reveals that more than 99% of the overall data variance summarizes in the first three principal components. As a result, on can expect that feature extraction, using for instance Fisher's Discriminant Analysis, will dramatically improve classification performances.

In Table 5.2, we present the overall classification scores obtained when applying the different classifiers described in the Section 5.3, with and without prior dimension reduction.

As expected, dimension reduction through FDA considerably improves the overall classification performances obtained with both k-NN and K-means classifiers. This suggests that Land Cover classes are better separated (w.r.t. Euclidean distance, at the core of theses

two classifiers) in the reduced eigenvectors hyper plane than in the original highly dimensional feature space. Moreover, results obtained with LDC not only indicate that data are reasonably linearly separable in the transformed space, but also that multivariate canonical distribution law is in good agreement with the normal hypothesis. Yet, it is worth noticing that to ensure linear separability of data with a 99% confidence interval, 11 eigenvectors were still needed (recall that initially we are dealing with reflectance vectors of dimension 13).

Table 5.2. Overall classification accuracies achieved with different classifiers and feature data sets.

Classifier	**Overall accuracy (%)**	
	Reflectance data	**FDA output data**
k-NN	65	76
K-means	31	67
SVM	73	78
LDC	-	77

Regarding SVM results, the influence of dimension reduction is not so drastic, in accordance with the well known fact that SVM is quite insensitive to the curse of dimensionality. Now, in terms of performances and computational costs, SVM clearly leaders when no dimension reduction is possible. Whenever FDA is affordable, then k-NN and LDC attain results very comparable to the ones achieved with SVM. However, the latter classifier is computationally more expensive, and LDC might be in that case a better choice.

Let us now split the global classification accuracy into generalization losses per land cover class (Table 5.3), considering the SVM classification of reflectance data linearly mapped by FDA. A closer view at the producer's and user's accuracy (respectively PA and UA) shows that "Broadleaved Closed Trees" (311) is definitely the less distinctive land cover class, frequently confounded with "Broadleaved Open Trees" (312) and "Mixed Closed Trees" (331). On the contrary, pixels of "Irrigated Herbaceous Crops" (22) and "Rice Crops" (23), two land cover classes that were very close according to mean spectral profiles of Figure 5.2, surprisingly almost never mix up. These results suggest that main classification puzzlement occurs among pixels of classes that contain the same land cover features but in different proportion arrangements. Land cover classes that consist of absolutely different surface elements, even with vegetation within their composition, are well separated. This demonstrates that the visual spectral similarity amongst different land cover classes does not imply their mixing, as attested by the overall accuracy results obtained with the exploited classifiers.

Table 5.3. User's and producer's accuracy per class. This repartition corresponds to a SVM classification of FDA transformed data based on reflectances measured at a single date (August 14, 2004).

Land cover class	**5**	**6**	**7**	**11**	**12**	**21**	**22**	**23**	**38**	**241**	**242**	**310**	**311**	**312**	**321**	**322**	**331**	**351**	**371**	**UA (%)**
5	39															1				98
6		25		1	3				7			1							3	63
7			40																	100
11		1	1	24	11		1				1		1							60
12		1		7	31				1											78
21						36				2									2	90
22							31	1		1	3			1			1	2		78
23							5	35												88
38		1			1				34			1				1		1	1	85
241						3		1		26	8			1					1	65
242										6	28		1	3				1	1	70
310												39	1							98
311									1	1		1	18	9		3	7			45
312													2	36		1	1			90
321															31	7	2			78
322									4			1	1		5	26		1	2	65
331							1						4	1	5		29			73
351									2						3			35		88
371						1			2		1	1			2		2		31	78
PA (%)	100	89	98	75	67	90	82	95	67	72	68	89	64	70	67	66	6	88	76	

5.5. Conclusion

The main goal of the current study was to evaluate the usefulness of using SVM for the characterization of numerous land cover classes with high spectral dimensional MERIS imagery data of Portugal continental. The achieved results suggest that SVM classifier outperforms k-NN, K-means and LDC classifiers in term of overall classification accuracy with both data sets – the full MERIS Level 2spectral imagery and the low-dimensional representation of such data derived through Fisher's discriminants. The level of classification accuracy achieved with SVM classifier is better than with the other classifiers when the whole spectral information is used. On the other hand, the k-NN and LDC performed a land cover classification with similar accuracy to SVM when a low-dimensional and rearranged representation of spectral data is used as input for classification task. However, unlike k-NN and LDC classifiers, SVM does not require a pre-processing technique to reduce data dimensionality in order to improve its classification accuracies, since they remained stable with full and low dimensional imagery data. The preliminary tests with the full spectral data

and compressed information with FDA showed that common classifiers had to face the problem of insufficient training data, which logically degraded the classification scores. In contrast, SVM proved to be insensitive to the dimensionality of input data, since classification results with MERIS data are improved, more stable and reliable than with other classifiers. From a methodological viewpoint, SVM are powerful and easy to tune learning systems that should rapidly enter the standard classifiers' toolbox used in remote sensing applications.

Regarding the particular land cover classes discrimination we think that MERIS imagery have an enormous capability to distinguish between them. The results proved that large mixing occurs only between land cover classes that contain the same land cover features but in different proportion arrangements. Classes with completely different configuration characteristics present a classification accuracy obtained with SVM classifier that is over 80%.

Future work will rely on the multi-temporal analysis of MERIS data, land cover classes redefinition, and data fusion appliance, since we consider that is possible to increase the overall classification accuracy obtained with MERIS images. We will also compare the classification accuracy obtained with MODIS data, a sensor with similar spatial and temporal resolutions to MERIS, but that acquires spectral information within short wavelength infrared.

5.6. Acknowledgements

This study was carried out in the framework of the project "User driven land cover characterisation for multi-scale environmental monitoring using multi-sensor earth observation data" funded by "Programa Dinamizador das Ciências e Tecnologias para o Espaço" from "Fundação para a Ciência e Tecnologia", and from the "Announcement of Opportunity for the Utilisation of ERS and ENVISAT Data" from European Space Agency (ESA).

5.7. References

Boser, B. E., Guyon, I. M., and Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In: Haussler, D., editor, Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory, 22-27 July, USA, pp. 144-152.

Clevers, J.G.P.W., Zurita Milla, R., Schaepman, M., and Bartholomeus, H. (2004). Using MERIS on ENVISAT for Land Cover Mapping. Proceedings of the 2004 Envisat & ERS Symposium, 6-10 September, 2004, Salzburg, Austria.

Cover, T. M., and Hart, P. E. (1967). Nearest Neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1): 21-27.

Cristianini, N., and Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press.

Curran, P.J. and Steele, C.M. (2005). MERIS: the re-branding of an ocean sensor. *International Journal of Remote Sensing*, 26(9): 1781-1798.

Di Gregorio, A., and Jansen, L.J.M. (2000). *Land Cover Classification System*. FAO, Rome, 179 pp.

Gessner, U., Gunther, K., P., and Maier, S. W. (2004). Landcover/land use map of Germany based on MERIS full resolution data. Proceedings of the 2004 Envisat & ERS Symposium, 6-10 September, 2004, Salzburg, Austria.

Gonçalves, P., Hugo Carrão, Andre Pinheiro and Mário Caetano (2005). Land cover classification with Support Vector Machine applied to MODIS imagery. Proceedings of the 25th EARSeL Symposium, June 6 - 11, 2005, Porto, Portugal.

Guyon I. and Elissee, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3: 1157-1182.

Haertel, Victor and David Landgrebe (1999). On the Classification of Classes with Nearly Equal Spectral Responses in Remote Sensing Hyperspectral Image Data. IEEE *Transactions on Geoscience and Remote Sensing*, l(37): 2374-2386.

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

Hughes, G. F. (1968). On the mean accuracy of statistical pattern recognizers. *IEEE Transactions on Information Theory*, 14: 55-63.

Jackson, Q. and David A. Landgrebe (2001). An Adaptive Classifier Design for High-Dimensional Data Analysis with a Limited Training Data Set. *IEEE Transactions on Geoscience and Remote Sensing*, 39(12): 2664-2679.

Jain, Anil K., Robert P.W. Diun, and Jianchange Mao (2000). Statistical Pattern Recognition: A Review. *IEEE Transactions on Pattern Analysis Intelligence*, 22(1): 4 –37.

Johnson, R. A. and Wichern, D. W. (1998). *Applied Multivariate Statistical Analysis* (4th ed.). Prentice Hall, Upper Saddle River, New Jersey.

Landgrebe, David (2002). Hyperspectral Image Data Analysis as a High Dimensional Signal Processing Problem, (Invited), *Special Issue of the IEEE Signal Processing Magazine*, 19(1): 17-28.

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

Maxwell, S.K., Hoffer R.M., and Chapman P. L. (2002). AVHRR channel selection for land cover classification. *International Journal of Remote Sensing*, 23(23): 5061-5073.

Mercier, G. and M. Lennon (2003). Support vector machines for hyperspectral image classification with spectral based kernels. Proceedings of IEEE Intl. Geoscience & Remote Sensing Symposium, IGARSS 2003 Conference, 21-25 July, Toulouse, France.

Mousavi, M., J., K. L. Butler-Purry, R. Gutierrez-Osuna and M. Najafi (2003). Classification of Load Change Transients and Incipient Abnormalities in Underground Cable Using Pattern Analysis Techniques. Proceedings of 2003 IEEE/PES Transmission and Distribution Conference, September 7-12, Dallas, TX.

Painho, M., and Caetano, M. (2005). *Cartografia de Ocupação do Solo, Portugal Continental, 1985-2000*. Instituto do Ambiente, Lisboa, Portugal.

Pal, M., and Mather, P. M. (2005). Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, 26(5): 1007-1011.

Price, K.P., X. Guo, and J.M. Stiles (2002). Optimal Landsat TM band combinations and vegetation indices for discrimination of six grassland types in eastern Kansas. *International Journal of Remote Sensing*, 23(23): 5031-5042.

Santos, T., Tenedório, J., Rocha, J., and Encarnação, S. (2005). SATSTAT: Exploratory Analysis of Envisat-MERIS Data for Land Cover Mapping of Portugal in 2003. Proceedings of 14th European Colloquium on Theoretical and Quantitative Geography, September 9-13, 2005, Tomar, Portugal.

Seeger, M. (2004). Gaussian processes for machine learning. *Interrnational Journal of Neural Systems*, 14 (2):1-38.

Shah, C. A., Watanachaturaporn, P., Arora, M. K., and Varshney, P. K. (2003). Some Recent Results on Hyperspectral Image Classification. In IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data. October 27-28, 2003, NASA Goddard Spaceflight center, Greenbelt, MD.

Suykens, J., Gestel, T., de Brabanter, J., de Moor, B., and Vandewalle, J. (2002). *Least Squares Support Vector Machines*. World Scientific.

Thomaz, C.E., R.Q. Feitosa, and A. Veiga (1999). Separate-Group Covariance Estimation with Insufficient Data for Object Recognition. Proceedings from the Fifth All-Ukrainian International Conference, Ukraine, pp. 21-24.

Vapnik, V. (1982). *Estimation of dependencies based on empirical data*. Springer Verlag, New York.

Vapnik, V. (1998). *Statistical Learning* Theory. Wiley.

Vieira, C. A. O., P. M. Mather, and P. Aplin (2001). Multitemporal classification of agricultural crops using the spectral-temporal response surface. In Analysis of Multi-Temporal Remote Sensing Images: Proceedings of the First International Workshop 2001, pp. 290-297.

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

6 MERIS BASED LAND COVER CLASSIFICATION WITH SELF-ORGANIZING MAPS: PRELIMINARY RESULTS

Carrão, H., Capão, L., Bação, F. and Caetano, M. (2006), MERIS Based Land Cover Classification With Self-Organizing Maps: Preliminary Results. In *Proceedings of the 2nd EARSeL SIG Workshop on Land Use & Land Cover* (unpaginated CD-ROM), 28-30 September 2006, Bonn, Germany.

Abstract

With the recent launch of MERIS, a wide range of new possibilities for the periodic land cover characterization at regional scale is available. This sensor offers a combination of innovative features, such as high spectral and temporal resolutions, wide geographical coverage and improved atmospheric correction. We believe that the exploitation of data obtained by this new sensor fills previous technological gaps, improving automatic land cover classes' discrimination. At the same time, the extra spectral information provided by MERIS can introduce some difficulties on land cover characterization with long-established classification techniques, e.g. k-Nearest Neighbour. In this paper we report the performance of artificial neural networks (ANNs) in the context of high spectral dimensional satellite image classification. The main goal of this research is to assess the potential of the Self-Organizing Maps (SOM) neural network to extract complex land cover type information from medium resolution satellite imagery. The study was carried out with MERIS Full Resolution data from 2004 for the continental Portuguese territory.

6.1. Introduction

Recently launched Earth Observation (EO) sensors, such as the MEdium Resolution Imaging Spectrometer (MERIS) and the Moderate Resolution Imaging Spectroradiometer (MODIS), exhibit enhanced spectral and temporal resolutions and offer new potentials and challenges to data analysis. Jackson and Landgrebe (i) state that the availability of a large number of spectral bands makes it possible to identify more detailed land cover classes with higher accuracy than would be possible with the data from earlier sensors, such as AVHRR. Quite as well, several studies have proved the advantages of performing a land cover classification based on multi-temporal satellite imagery data (e.g. ii, iii, iv, v). However, the exploitation of spectral and temporal characteristics of MERIS imagery data, for the improvement of land cover spatial and thematic characterization, is still of confidential use in remote sensing

community. Presently, few research studies and operational programs are focused on the analysis of MERIS data for land cover cartography production (e.g. vi, vii, viii).

The introduction of these new generation sensors resulted in a large increase of remote sensing data volumes available for land cover classification purposes. Still, high dimensional remote sensing imagery provides a challenge to traditional classification methods based on statistical assumptions. Artificial Neural Networks (ANNs) may represent a viable alternative approach to land cover mapping using high dimensional remote sensing imagery. ANNs require no assumption regarding the statistical distribution of the input pattern classes, are relatively noise tolerant and entail a massively parallel structure (ix). Neural network models have two important properties: the ability to "learn" from input data and to generalize and predict unseen patterns based on the data source, rather than on any particular *a priori* model (x).

Although ANNs are routinely applied to multispectral and SAR imagery classification, the number of studies that utilize ANNs for high spectral dimensional image analysis are limited (ix). These authors also refer that existing studies focused primarily upon supervised approaches, while there is little precedent specifically for the utilization of the unsupervised ANNs in remotely sensed image classification, especially with respect to hyperspectral analysis. However, unsupervised ANN paradigms based on competitive learning have simple architectures and faster learning algorithms than supervised Multi-Layer Perceptrons (MLPs) (xi). Some researches have shown that the unsupervised Self-Organizing Maps (SOM) technique possesses a speedier training rate and higher classification accuracy when classifying remotely sensed data then the supervised backpropagation neural network (xii). The Self-Organizing Map, developed by Kohonen (xiii), owes its popularity to its capability of presenting on the output layer, typically arranged in a two-dimensional grid, a very intuitive description of the similarity among groups in the input space. The SOM was not originally intended for pattern classification, but a subsequent calibration phase can always be applied to produce labels for the clusters resulting from its application. Once an unsupervised ordering step is performed, the feature map is trained as a classifier using a feature vector training set with known classifications (ix). Although several supervised extensions of SOM have already been implemented for classification, the most widely used has been the Linear Vector Quantization (LVQ) procedure (xiv).

In this paper we exploit the SOM algorithm for land cover classification with MERIS high spectral dimensional imagery. This is a preliminary study in the framework of an ongoing research work that aims at developing a systematic classification methodology for the

regular production of land cover cartography from medium spatial resolution satellite imagery in Portugal. To completely evaluate the ability of the high spectral resolution MERIS data for land cover characterization, we look for the separation of 19 land cover classes. Some results comparing the SOM performance with the traditional k-Nearest Neighbour method are also presented. We experimentally evaluate the performance of each classifier on samples collected from a summer MERIS image of Portugal Continental.

6.2. Study area and data set

The study area is the entire Portuguese mainland territory. Portugal is in a transition zone featuring diverse landscapes representing both Mediterranean and Atlantic climate environments. This landscape heterogeneity allows for the extrapolation of the developed methodologies and of the obtained results to many other regions of the world.

In this study we exploit the Level 2 Full Resolution MERIS imagery. These data consists of calibrated surface reflectances in 13 groups of wavelengths (original bands 11 and 15 were removed from this product since they address O_2 content and water vapor) that are atmospherically corrected for Rayleigh scattering, ozone, water vapor absorption and aerosol content. We make use of a single image that was acquired by 14 August of 2004. This date was chosen due to several criteria. First, and according to (ii), August is the period of the year that allows for a better separation among general land cover classes in this region of transitional climate; secondly, this image is the one that presents the least cloud coverage among MERIS August images available for Portugal for the year of 2004.

The MERIS original data was exported from raw format and converted to GeoTIFF using the BEAM VISAT 3.4 ® software application. Also, in order to combine these data with already existing auxiliary information it was necessary to project the image into a proper map projection (Hayford-Gauss, Datum Lisboa). This was easily accomplished since MERIS Level 2 imagery are already geolocated.

To help the selection of samples for algorithm training, validation and testing we made use of existent land cover databases and high spatial resolution Earth Observation (EO) data, namely: the CORINE Land Cover map for 2000 (CLC2000) (xv) – land cover map with 44 classes derived from visual interpretation of Landsat-7 ETM+ images acquired in the summer of 2000; the National cover of SPOT High Resolution Geometric (HRG) images acquired in 2003; the National cover of orthorectified colour infrared aerial photography for 1995; and the National Forest Inventory for 1995.

6.3. Methodology

6.3.1. Land cover nomenclature

In this study we used the land cover nomenclature developed by the Remote Sensing Unit of the Geographic Portuguese Institute (IGP) (Table 6.1). The nomenclature classes were defined through the Land Cover Classification System (LCCS) from Food and Agriculture Organization (FAO) (xvi). The rationale behind the development of the nomenclature was two-fold: (1) a nomenclature that is well adapted to the type of landscapes existent in regions with characteristics similar to the Portuguese mainland, and (2) a nomenclature that is compatible with established ones (e.g., CORINE Land cover, Global Land cover and the International Geosphere-Biosphere Programme nomenclatures) in order to make possible map comparison.

Table 6.1. Land cover classes and respective acronym.

Land Cover Class	**Acronym**
Continuous Artificial Areas	11
Discontinuous Artificial Areas	12
Rain fed Herbaceous Crops	21
Irrigated Herbaceous Crops	22
Rice Crops	23
Permanent Evergreen Crops (Trees or Shrubs)	241
Permanent Deciduous Crops (Trees or Shrubs)	242
Broadleaved Closed Trees	311
Broadleaved Open Trees	312
Needleleaved Closed Trees	321
Needleleaved Open Trees	322
Mixed Closed Trees	331
Shrubland	35
Natural grassland	37
Sparse Vegetation	38
Recently Burnt (Trees or Shrubs)	310
Permanent Wetlands	5
Barren	6
Water Bodies	7

6.3.2. Land cover sampling

In order to test the several classification approaches using the defined nomenclature, we carried out a collection of samples for the 19 nomenclature classes for the year of 2004. The sampling process was performed by visual interpretation of the high spatial resolution EO data and land cover databases described in the Study Area and Dataset section.

Each sample represents a specific land cover class covering at least 90% of the area enclosed by a MERIS 300-by-300 m pixel (same as the nominal resolution of used satellite images). For each class we collected 40 samples distributed all over the mainland territory in order to obtain a broad representation of within classes' spectral differences.

6.3.3. Classification phase

Kohonen's Self-Organizing Map (SOM)

Kohonen's self-organizing map is one of the most interesting of all the competitive neural nets (xvii). A SOM is an unsupervised and nonparametric neural network approach, which means that no human intervention is needed during the learning and that little need to be known about the characteristics of the input data. The SOM is remotely inspired in the way in which various human sensory impressions are neurologically mapped into the brain, so as spatial or other relations among stimulus correspond to spatial relations among the neurons (xviii). As a neurospatial classifier, the SOM offers a number of attractions, namely the simple design and ability to handle immense complexity (xvii). A SOM consists usually of a two-dimensional array of neurons fully connected with the input vector, arranged on a squared or hexagonal lattice, such that there are neighborhood relationships among the neurons (xi). The latter "architecture" (hexagonal) is generally preferred, because it provides a better coverage of the two-dimensional space and an easier visualization of the input space structure (xi, xvii).

The training procedure of SOM consists of finding the neuron with weights closest to the input data vector and declaring that neuron as the winning neuron. As a result, the closest neuron to input vector is updated, as well as all neighbour neurons within a predefined in the output space. Several functions can be used to determine the neighbors' movements during training, while the most commonly adopted is the Gaussian (xi). By assigning each input vector to the nearest neuron, the SOM is able to organize the output space into regions, where neighboring neurons have similar characteristics. In a high dimensional input space, Kohonen maps can be interpreted as nonlinear projectors of the input data onto a two-dimensional space that takes into account the data probability density and preserves the original topology of the input patterns (xi). That is, if two weight vectors are near each other in the input space, the corresponding neurons will also be close in the output space, and vice versa.

The SOM clustering of MERIS data was carried out using the MATLAB SOM Toolbox (xix).The Kohonen's SOM algorithm that is implemented in the SOM Toolbox can be

described as follows: consider that there are M neurons in the competitive layer and every neuron j has a n-dimensional weight vector, $w_j=[w_{j1}, \ldots, w_{jn}]$. At each training step *t*, a sample data vector $x(t) = [x(t)_1, \ldots, x(t)_n]$ is randomly chosen from the training set (considering a sequential training approach). The Euclidean distance is used to determine the winning neuron. The winning neuron or best-matching unit (BMU), denoted here by c, is the neuron with the weight vector closest to $x(t)$:

$$a_j = \left(\sum_{i=1}^{n} (x(t)_i - w_{ji})^2 \right)^{0.5} = \left\| x(t) - w_j \right\|, j \in \{1.....M\} \quad (6.1)$$

if $a_c = \min\{a_1, a_2, ..., a_M\}$, c is the winning neuron.

Update weights for *c* and neighbor neurons within neighborhood *Nc(t)* can be written as:

$$w_{ji}(t+1) = w_{ji}(t) + \alpha(t)[x(t)_i - w_{ji}(t)] \text{ if } j \in N_c(t) \quad (6.2)$$

and

$$w_{ji}(t+1) = w_{ji}(t) \text{ if } j \notin N_c(t) \quad (6.3)$$

where *Nc(t)* is the set of neighborhood neurons of the winning neuron at time *t*, $\alpha(t)$ is the learning rate, and its initial value is set as $0<\alpha(t)<1$. Both learning rate $\alpha(t)$ and neighborhood (*Nc(t)*) radius decreased monotonically with time, in order to assure convergence. After finding the winning neuron, the weight vector of the neuron in the competitive layer is updated. During the update procedure, the winning neuron is moved closer to the sample vector in the input space. The topological neighbours of the winning neuron are updated in a similar way. This update procedure stretches the winning neuron and its topological neighbours towards the sample vector. Neighbouring neurons are pulled in the same direction, and thus weight vectors of neighbouring neurons resemble each other for adopting competitive learning rules and neighbourhood (the topological neighbourhood relationships). Neurons with the same properties in the competitive layer would be close each other, and the others would be far away when the clustering process is finished. Consequently, different types of clusters of neurons were generated on the competitive layer. At the end of the training phase each neuron is labeled to a corresponding land cover class according to a majority voting principle (by taking into account the land cover classes of the closest training samples data vectors).

The structure of the SOM network and respective parameters was object of several experiments in order to attain the best overall land cover classification accuracy with the MERIS data. Considering the literature reviewed, the topology of the SOM was always set

up as an hexagonal lattice. In the same way, a Gaussian neighbourhood function with linear descending rate was used in all configurations. An initial learning rate $\alpha(0)$ = 0.7 was established. An initial learning rate value between 0.1 and 0.7 must always be set (xi).

SOM generalization performance evaluation. The application of the SOM to automatic land cover classification with MERIS data followed three main steps: training (also called learning), validation, and independent performance assessment (or testing). Thus, the 40 land cover samples that were collected for each of the 19 classes were randomly divided in three different sets, in the following way: training (70%), validation (15%) and independent testing (15%). The goal in training the network is to find a structure that performs good classifications for both the training data and for new data, avoiding data overfitting. Therefore, when a training step finishes, the validation component is presented to the network and the generalized classification error is evaluated. The network having the lowest validation error will be used for land cover classification, and the overall performance of the selected model is evaluated using the independent testing component.

k-Nearest Neighbor (k-NN)

k-NN is an instance-based learning algorithm (xx) based on a distance function (e.g. Euclidian distance) between pairs of observations. The algorithm is quite simple: given a feature test vector, the system finds the k-nearest neighbours among the training vectors, and uses the labels of these k neighbours to determine the class of the unknown sample. Using the training set as labeled prototype vectors, we evaluate the classification results of the validation set considering 1 to 10 nearest neighbours. The k-NN configuration with better overall accuracy results was chosen to use for the evaluation of SOM classification results.

6.4. Results and discussion

Trial-and-error experiments with the training and validation sets were conducted to empirically determine the best architecture and parameters for the SOM network. Regarding network size, trials were conducted first with small networks (e.g., as small as 25 neurons), which resulted in some unrepresented classes in the final labeled network. Consequently, the overall classification results were extremely small in those situations, which did not exceed the 30% of agreement. Classes as “Mixed Closed Trees” (331), or “Rain fed Herbaceous Crops” (331), were not presented in the final map, being confounded with other several land cover classes of spectral similarities. Larger competitive layers were thus used to map land cover classes at a finer resolution. However, the larger 2-D SOM network that could be constructed, given computer memory constrains, was approximately 38x38 (1444 neurons).

The results tend to improved as the size of the network increases, thus larger networks lead to improved overall classification accuracy.

The parameters used for SOM network simulations varied according to the size of the hexagonal lattice used. The initial learning rate was always 0.7, and a reciprocally decreasing function was used to adjust learning rate from first to last training epoch. A Gaussian neighborhood function was always used to adjust the weights of all of the neurons in the neighborhood of the winning neuron, but larger radiuses were used in larger networks. A linearly decreasing function was used to adjust the network radius during training. Regarding the training length, bigger networks have longer learning periods. The parameters values and functions were selected according to the suggestions purposed by (xxi).

The overall classification accuracy was evaluated with the test set for the network with the smaller classification error obtained with the validation set (each validation sample was classified in the land cover class of the nearest neuron after the training step). This condition was attained with a lattice with 1444 neurons, and the overall classification accuracy was 68% with the test set. Although these results are good and more effective than those attained with k-NN with a neighborhood of three (65%), they do not outperform the 73% overall classification accuracy achieved by (xxii) with the Support Vector Machine (SVM) for the same data set. We believe that these empirical results can be augmented, because the most suitable set of parameters and architectures of SOM are difficult to tune and there are still several unexplored possibilities to optimize the SOM's results.

A closer view at the producer's (PA) and user's (UA) accuracy (Table 6.2) shows that "Sparse Vegetation" (38) is definitely the less distinctive land cover class, frequently confounded with "Needleleaved Open Trees" (322). It is also interesting to look at the confusion among "Barren" (6) and "Natural grassland" (371), that is certainly due to the similar spectral characteristics of these two land cover classes on summer. Agricultural classes present excellent classification accuracies, except the Permanent Crops (241 and 242) which tend to be mixed. On the contrary, pixels of "Water bodies" (7) and "Continuous Artificial Areas" (11) never mix up. These results suggest that main classification confusion occurs among pixels of classes that contain the same land cover features but in arrangements with different proportions. Land cover classes that consist of absolutely different surface elements, even with vegetation within their composition, are well separated (e.g., "Shrubland" (351) and "Broadleaved Closed Trees" (311)). This demonstrates that noticeable spectral similarity amongst different land cover classes does not imply their mixing, as attested by the overall accuracy results obtained with the exploited classifiers.

Table 6.2. User's and producer's accuracy per class. This repartition corresponds to the classification of test samples performed with the SOM network with 1444 neurons.

Land cover class	**5**	**6**	**7**	**11**	**12**	**21**	**22**	**23**	**38**	**241**	**242**	**310**	**311**	**312**	**321**	**322**	**331**	**351**	**371**	**UA (%)**
5	5																			100
6		3							1											75
7	1		6																	86
11				6	1															86
12		1			5															83
21						5													1	83
22							5				1									83
23							1	5												83
38									1				1			1		1	1	20
241						1				4	2									57
242										2	2									50
310												4							1	80
311													4	4						50
312														2						100
321															3	1				75
322									4						2	4				40
331													1		1		6	1		67
351								1										4		80
371		2									1	2							3	38
PA (%)	83	50	100	100	83	83	83	83	17	67	33	67	67	33	50	67	100	67	50	

6.5. Conclusion

The results show that SOM is more effective than k-NN for land cover classification with MERIS data. The results also show that, as expected, larger networks will produce better results. The difference accuracy, between the SOM and the k-NN, is small (3%) which may be considered disappointing, nevertheless we think that there are several unexplored avenues to optimize the SOM's results. From a methodological viewpoint the SOM has proved to be a powerful algorithm, but it requires an experienced user because of the number of parameters that need to be tuned. Thus, the determination of the most suitable set of parameters and architecture constitutes a future research problem. Another aspect that needs further research is related with the possibility of using a number of SOM instead of using only one SOM for the classification of all classes. This will create the opportunity to develop "specialized" SOM's which are particularly good in the identification of certain classes.

Regarding the particular land cover classes discrimination, we think that MERIS imagery has an enormous capability to make a good quality distinction among them. The results proved that large mixing occurs only between land cover classes that contain the same land cover features but in different proportion arrangements. Classes with completely different configuration characteristics present a classification accuracy obtained with the SOM neural network that is over 80%.

Future work will rely on the multi-temporal analysis of MERIS data, land cover classes redefinition, and appliance of supervised artificial neural network methods, since we consider that is possible to increase the overall classification accuracy obtained with MERIS images. In the future, a supervised extension of the competitive learning process, like the Learning Vector Quantization (LVQ), should also be exploited as a technique for automatic land cover classification.

6.6. Acknowledgements

This study was carried out in the framework of the project "LANDEO - User driven land cover characterisation for multi-scale environmental monitoring using multi-sensor earth observation data (PDCTE/MGS/49969/2003)" funded by "Programa Dinamizador das Ciências e Tecnologias para o Espaço" from "Fundação para a Ciência e Tecnologia", and from the "Announcement of Opportunity for the Utilisation of ERS and ENVISAT Data" from European Space Agency (ESA). Research by Hugo Carrão was founded by the "Fundação para a Ciência e Tecnologia" (SFRH/BD/18447/2004).

6.7. References

i. Jackson Q & D Landgrebe, 2001. An Adaptive Classifier Design for High-Dimensional Data Analysis with a Limited Training Data Set. IEEE Transactions on Geoscience and Remote Sensing, 39: 2664-2679.

ii. Gonçalves P, H Carrão, A Pinheiro & M Caetano, 2006. Land cover classification with Support Vector Machine Applied to MODIS imagery. In: Global Developments in Environmental Earth Observation from Space, edited by A Marçal (Rotterdam: Millpress), 517-526.

iii. Price K P, X Guo & J M Stiles, 2002. Optimal Landsat TM band combinations and vegetation indices for discrimination of six grassland types in eastern Kansas. International Journal of Remote Sensing, 23: 5031-5042.

iv. Maxwell S K, R M Hoffer & P L Chapman, 2002. AVHRR channel selection for land cover classification. International Journal of Remote Sensing, 23: 5061-5073.

v. Vieira C A O, P M Mather & P Aplin, 2001. Multitemporal classification of agricultural crops using the spectral-temporal response surface. In: Analysis of Multi-Temporal Remote Sensing Images – Series in Remote Sensing (Vol. 2), edited by L Bruzzone & P Smits (Singapore: World Scientific), 290-297.

vi. Clevers J G P W, R Zurita Milla, M Schaepman & H Bartholomeus, 2004. Using MERIS on ENVISAT for Land Cover Mapping. In Proceedings of the 2004 Envisat & ERS Symposium, 6-10 September, Salzburg, Austria.

vii. Gessner U, K P Gunther & S W Maier, 2004. Landcover/land use map of Germany based on MERIS full resolution data. In Proceedings of the 2004 Envisat & ERS Symposium, 6-10 September, Salzburg, Austria.

viii. Santos T, J Tenedório, J Rocha & S Encarnação, 2005. SATSTAT: Exploratory Analysis of Envisat-MERIS Data for Land Cover Mapping of Portugal in 2003. In Proceedings of 14th European Colloquium on Theoretical and Quantitative Geography, 9-13 September, Tomar, Portugal.

ix. Filippi A M & J R Jensen, 2006. Fuzzy Learning Vector Quantization for Hyperspectral Coastal Vegetation classification. Remote Sensing of Environment, 100: 512-530.

x. Seto K & W Liu, 2003. Comparing ARTMAP neural network with the maximum-likelihood classifier for detecting urban change. Photogrammetric Engineering & Remote Sensing, 69: 981-990.

xi. Berthold M & D J Hand, 2003. Intelligent Data Analysis (Berlin: Springer Publications, 2nd ed.) 514 pp.

xii. Ji C Y, 2000. Land-use classification of remotely sensed data using Kohonen Selforganizing feature map neural networks. Photogrammetric Engineering and Remote Sensing, 66: 1451-1460.

xiii. Kohonen T, 1984. Self-organization and associative memory (Springer Verlag) 501 pp.

xiv. Kohonen T, 1995. Self-organizing maps. (Springer Verlag) 501 pp.

xv. Painho M & M Caetano, 2005. Cartografia de Ocupação do Solo, Portugal Continental, 1985-2000. Instituto do Ambiente, Lisboa, Portugal.

xvi. Di Gregório A & L J M Jansen, 2000. Land Cover Classification System (Rome: FAO), 179 pp.

xvii. Openshaw S & C Openshaw, 1997. Artificial Intelligence in Geography (Chichester: John Wiley & Sons Ltd) 329 pp.

xviii. Sarle W S, 1997. Neural Network FAQ, part 1 of 7: Introduction, periodic posting to the Usenet newsgroup comparing neural-nets. Retrieved 25 April 2006 from: ftp://ftp.sas.com/pub/neural/FAQ.html

xix. Vesanto J, J Himberg, E Alhoniemi & J Parhankangas, 1999. Self-organizing map in matlab: the SOM toolbox. In Proceedings of the Matlab DSP Conference 1999, Espoo, Finland, 35 - 40.

xx. Cover T M & P E Hart, 1967. Nearest Neighbor pattern classification. IEEE Transactions on Information Theory, 13: 21-27.

xxi. Vesanto J, J Himberg, E Alhoniemi & J Parhankangas, 2000. SOM Toolbox for Matlab 5. Technical Report A57, Helsinki University of Technology, Neural Networks Research Centre, Espoo, Finland.

xxii. Carrão H, P Gonçalves & M Caetano, 2006. MERIS Based Land Cover Characterization: A Comparative Study. In Proceedings of the ASPRS 2006 Annual Conference - Prospecting for Geospatial Information Integration, 1-5 May, Reno, USA.

7 SAMPLE DESIGN AND ANALYSIS FOR THEMATIC MAP ACCURACY ASSESSMENT: AN APPROACH BASED ON DOMAIN ESTIMATION FOR THE VALIDATION OF LAND COVER PRODUCTS

Carrão, H., Caetano, M. and Coelho, P. (2007), Sample Design and Analysis for Thematic Map Accuracy Assessment: An Approach Based on Domain Estimation for the Validation of Land Cover Products. In *Proceedings of the 32nd International Symposium on Remote Sensing of Environment* (unpaginated CD-ROM), 25-29 June 2007, San Jose, Costa Rica.

Abstract

Remote sensing data has been extensively used to generate land cover and land use maps for a variety of purposes using automatic classification approaches. However, while major efforts have been made in the development and description of automatic techniques for land cover classification, the survey sampling for the accuracy assessment of land cover maps evaluation is either insufficient or incorrectly reported. In this paper we review the basic components that constitute a statistically rigorous accuracy assessment for land cover maps evaluation. Simultaneously, we focus on the adequate estimation of overall, user's and producer's accuracies under a stratified sampling design and introduce an unexploited approach for their calculation based on domain estimation.

Keywords: sample design, accuracy assessment, land cover products, domain estimation.

7.1. Introduction

Land cover products derived from satellite image data became very popular in forestry, geology, landscape analysis and management, giving real time information about the spatial distribution of natural features and their changing patterns. However, before being used in scientific investigations and policy decisions, these thematic maps should be subject to a statistically rigorous accuracy assessment. Accuracy assessment estimates general and specific qualities of the derived land cover products so that users may evaluate their utility for specific applications (Stehman and Czaplewski, 1998; Stehman, 2001). Users will not and should not take a map at face value without some associated estimate of error (Card, 1982). Thus, accuracy assessment should be an integral component of any mapping project based on remote sensing data. A census is, of course, impractical to determine map error, so the key component to any effective accuracy assessment is the employment of a practical and

statistically sound sampling survey to characterize the accuracy of common and rare classes of map products. Indeed, a more detailed report of map accuracy in the form of individual land cover category accuracies is demanded (Zhu *et al.* 2000).The categories could, and frequently do, exhibit drastically differing accuracies, and yet combine for equivalent or similar overall accuracies. Individual category estimators of accuracy are, therefore, needed to completely assess the validity of the map for a specific application.

At national and regional level it is difficult to focus whole landscape diversity in a map comprising a set of restrict and predefined land cover classes, because there is a low consistency in landscape characteristics at such scale. Thus, several automatic classification approaches are usually tested and numerous land cover products need to be evaluated until an acceptable map is available to be applied by end users. Individual category estimators for each map can only be derived through stratified random sampling, which guarantees a minimum sample size in each stratum or category (Stehman, 1996). However, successive sample collections for the validation of each map product are timely and economically expensive. Therefore, there is a need to reuse a sample that was already collected even by using the land cover classes of a previous map as strata. Still, Stehman (1996) states that this can be impractical because stratifying by mapped land cover locks the assessment into the map version used to form the strata.

The goal of this paper is three-fold: 1) to review the basic components that constitute a statistically sound survey sampling for land cover maps accuracy assessment; 2) to present the adequate estimations for overall and per class accuracies in the framework of stratified sampling; and 3) to introduce an unexploited method on domain estimation – based on the works developed by Cochran (1977, Chapter 5) and Särndal *et al.* (1992, Chapter 10) – for more precise overall and per class accuracies estimations of revised maps by reusing samples collected within stratified sampling.

7.2. Survey sampling

In designing a survey sampling for accuracy assessment, one must keep in mind the acceptable level of accuracy, the reference sample, the sample selection strategy and the sample size required to estimate overall and per class accuracies associated with some land cover map (Dicks and Lo, 1990). The quantification of mapping errors within specified confidence limits requires that appropriate sampling techniques and sample sizes are selected adequately. The acceptance/rejection criteria for judging any survey sampling for accuracy assessment of land cover maps have been summarized by Ginevan (1979) as follows: 1) there should be a low probability of accepting a map of low accuracy; 2) there should be a

high probability of accepting a map of high accuracy; and 3) a minimum number of ground data sample points should be required.

7.2.1. Target accuracy

Any accuracy assessment must begin with the definition of the acceptable level of accuracy. The required level of accuracy of the final product should depend not only on the feasible results of some automatic classification approach and input data characteristics, but also on its intended application. If the user requests a map with some level of accuracy, then it is imperative the producer to generate it or it will be of no use. If a map is expected to be less than adequate for a given application, resources should be expended on improving the classification accuracy to the desired level (Wulder *et al*. 2006).

7.2.2. Reference sample

A sample observation serves as the basic element of comparison between the land cover map classification and the "reference" classification. Stehman and Czaplewski (2003) suggest that "reference classification" should be used to avoid the phrase "true classification", because the true land cover of an observation can very upon the choice of definitions and measurement protocols. A sample observation can be a point or an area and its various forms and relative advantages and disadvantages were examined in detail by Stehman and Czaplewski (1998). In the context of land cover products derived from pixel-based classification of satellite imagery, the sample observation is typically the minimum mapping unit of the map, i.e. a pixel area, which is the same as the source imagery (Whickam *et al*., 2004).

7.2.3. Probability sampling design

The sampling design is the set of rules used for selecting the reference sample (Stehman, 1999). A probability sampling design is one in which the inclusion probabilities are known for all observations in the sample and are non-zero for all elements in the population (Stehman, 2001). A statistically rigorous design contributes to a scientifically defensible accuracy assessment, and such designs must be used because of their objectivity (Stehman and Czaplewski, 1998). If this is not the case, the inference becomes entirely dependent on assumptions and, as such, will be very difficult to defend in a confrontational setting (Stehman, 2001). For most practical problems encountered in accuracy assessment of land cover maps, a statistically sound sampling procedure exists (e.g., simple random, stratified random, cluster, and systematic sampling) – for a complete description of common

probability sampling designs see, e.g. Stehman and Czaplewski (1998), Stehman (1999) and Congalton and Green (1999).

Stratified sampling must be applied when there is a need to ensure a minimum sample size in each stratum to derive accuracy estimates for all land cover classes presented in the map (Stehman, 1999). Moreover, stratifying by mapped land cover classes may be used to achieve more precise estimates than would be obtained from simple random or systematic sampling without stratification (Stehman, 2001).

7.2.4. Sample size

The sample size must be selected with care, because it should be sufficient to provide a representative and meaningful basis for accuracy assessment (Hay, 1979). In general, the larger the sample size the greater the confidence one can have in assessments based on that sample (Dicks and Lo, 1990). Several authors have suggested methods and guidelines to estimate the appropriate sample size. The goal is to provide accuracy estimations preferably with small variance estimates at some high level of confidence. Along years, main authors have used an equation based on the normal approximation to the binomial distribution to estimate the required sample size for the accuracy assessment of land cover maps (e.g. Hay, 1979; Dicks and Lo 1990). Congalton and Green (1999) state that this approach is statistically sound for estimating the sample size needed to estimate the overall accuracy of a classification or the accuracy of a single land cover category.

7.3. Analysis and estimation

The main goal of a survey sampling for land cover cartography accuracy assessment is to provide information to estimate the overall proportion of area correctly classified in the map and per cover category, as well as the confidence intervals for those estimates. For that reason, the map and reference classifications determined for each sample observation are usually displayed in a “contingency table” or “error matrix” in the form of a cross-tabulation of map category versus reference category. Each cell entry represents the number of sample observations whose map category is the row label and whose reference category is the column label. This table gives a visual overview of the map accuracy results and introduces a first estimate of two kinds of per class accuracies inherent in the classification system: user’s and producer’s accuracies for each cover type (Card, 1982). The first is the proportion of a mapped land cover class that was classified in accordance with the reference, while the second is the reference proportion of that class that was correctly mapped. Once the sample observations are compiled in this kind of table, we can make unbiased estimates of those accuracies of interest. Most land cover cartography accuracy estimators, e.g. user’s,

producer's and overall accuracies, can be formulated from marginal totals of the contingency table. Therefore, they can be simply estimated by the Horvitz-Thomson (HT) estimator or as a function of HT estimators. In addition, estimators' precision can be improved at the analysis stage by post-stratified estimation (Card, 1982). Post-stratified estimators, which demand the integration of auxiliary variables in the estimation, are always available for land cover maps accuracy assessment because category areas are known in the map and can be used to increase estimator precision. This estimation technique increases the complexity of the analysis, but usually incurs no additional sampling cost (Stehman and Czaplewski, 1998; Stehman, 2001). Weighting estimation of overall and per class accuracies by mapped land cover classes' areas is well illustrated by the formulas for stratified sampling in land cover cartography assessment presented by Card (1982).

To reduce sampling efforts, the sample collected through a stratified random sampling design for some basis land cover map must be reused in the estimation of overall and per class accuracies of revised maps. The land cover classes in the revised maps can be considered as subpopulations or domains represented in all strata (land cover classes) of the basis map. The accuracy estimators for newer maps present some complications, as stated by Cochran (1977, Chapter 5), because the number of observations in the sample that fall into each domain is a random variable. Since observations in the sample are assumed to be observed without error, the only randomness in the model comes from the probabilities of selection which depend on the sampling design. We present the formulas for the overall and per class accuracies estimators, and their respective variance estimates, for newer maps following a domain estimation method.

7.3.1. Description of estimators

The accuracy estimators cannot be correctly derived from a contingency table unless the sampling survey which was followed in collecting the sample is known. At least, intelligent inferences to the underlying population cannot be made (Card, 1982). A proper procedure for the estimation of the population totals, variances and confidence limits, to be used under a stratified random sampling approach, was presented by Cochran (1977, Chapter 5), and designed for map accuracy assessment by Card (1982). Card (1982) points out that, for the stratified sampling case, the overall proportion of correctly classified individuals in the map, given the "reference" land cover categories, should not be simply estimated by the diagonal entry divided by the row sum of the contingency table, because of the bias introduced by possible differential sampling rates within map categories. Therefore, the overall and per class accuracy estimations that will be used to evaluate the land cover map should include

the known areas (marginal distributions - N_h) of each map category (h) to improve the estimation of the proportion of correctly mapped individuals. In practice, suppose that a sample of n observations is located on the map and the random variable G, i.e. the "reference" land cover category, is determined for each sample observation in each stratum h (h = 1, 2,…, H map categories). Both land cover categories assembled for sample observations can be tabulated in a two-way square contingency table, as shown in Table 7.1, where n_{hg} is the number of sample observations whose map category is h and whose reference category is g.

Table 7.1. Sample error matrix.

Land cover map category (Stratum) \ Reference land cover category	G_1	G_2	⋯	G_g	*RowTotal*
H_1	n_{11}	n_{12}	⋯	n_{1g}	n_{1+}
H_2	n_{21}	n_{22}	⋯	n_{2g}	n_{2+}
⋮	⋮	⋮	⋱	⋮	⋮
H_h	n_{h1}	n_{h2}	⋯	n_{hg}	n_{h+}
ColumnTotal	n_{+1}	n_{+2}	⋯	n_{+g}	n

In a stratified random sampling, the sample sizes per stratum, n_h , are specified in advance of the sample collection according to the survey sampling specifications, and each stratum is sampled independently of the other strata. Thus, for a stratified sampling design, in which the strata are the map categories, the row totals, n_{h+}, are fixed, but the column totals, n_{+g}, are random variables that depend on the observed sample. In addition, and considering that the thematic map consists of N pixels, the pixel totals for each land cover category in the map, N_h, are known, but the pixel totals for each reference land cover category, N_g, are unknown. Similarly, all entries, N_{hg}, within an hypothetic population error matrix are also unknown.

We finish this paper by presenting the details about the unbiased derivation of the user's, producer's and overall accuracy estimators and respective variance, under a stratified random sampling survey. These estimates followed the results that Card (1982) published in the framework of accuracy assessment of land cover maps. Note that these derivations are an extension of the general formulas for stratified sampling estimation presented earlier by Cochran (1977, Chapter 5). Suppose that:

N – number of pixels in the map;

N_h – number of pixels in the map category h;

N_{h1} – number of pixels in the map category h that correspond to same reference category;

N_g – number of pixels in the reference category *g*;

n_h – number of pixels in the sample collected in map category *h*;

n_{hg} – number of pixels in the sample collected in map category *h* that intersect reference category *g*;

n_{h1} – number of pixels in the sample collected in map category *h* that correspond to same reference category.

The overall accuracy estimation for the map, i.e. the proportion of correctly classified pixels in the map, is

$$\hat{P}_c = \frac{1}{N}\sum_{h=1}^{H}\frac{N_h}{n_h}n_{h1} \tag{7.1}$$

and the respective estimated variance

$$\hat{V}(\hat{P}_c) = \sum_{h=1}^{H}\left(\frac{N_h}{N}\right)^2 \frac{N_h - n_h}{N_h n_h}\left[\frac{n_{h1}}{n_h}\left(1-\frac{n_{h1}}{n_h}\right)\right] \tag{7.2}$$

The user's accuracy estimation for each land cover class in the map is

$$\hat{P}_{h,c} = \frac{n_{h1}}{n_h} \tag{7.3}$$

and the respective estimated variance

$$\hat{V}(\hat{P}_{h,c}) = \left(1-\frac{n_h}{N_h}\right)\frac{1}{n_h - 1}\left[\frac{n_{h1}}{n_h}\left(1-\frac{n_{h1}}{n_h}\right)\right] \tag{7.4}$$

The estimation of producer's accuracy under a stratified random sampling survey must be calculated using an unbiased estimation of the population unknown quantities N_{h1} and N_g. Their unbiased estimators are:

$$\hat{N}_{h1} = \frac{N_h}{n_h}n_{h1} \tag{7.5}$$

$$\hat{N}_g = \sum_{h=1}^{H}\frac{N_h}{n_h}n_{hg} \tag{7.6}$$

Then, the producer's accuracy can be viewed as a ratio of these two estimators as follows:

$$\hat{P}_{g,c} = \frac{\hat{N}_{h1}}{\hat{N}_g} \tag{7.7}$$

and the respective estimated Mean Square Error (MSE)

$$\hat{MSE}(\hat{P}_{g,c}) = \left(\frac{N}{\hat{N}_g}\right)^2 \sum_{h \neq g}\left[\left(\frac{N_h}{N}\right)^2 \left(\frac{N_h - n_h}{N_h n_h}\right)\frac{n_{hg}}{n_h}\left(1 - \frac{n_{hg}}{n_h}\right)\hat{P}_{g,c}^2\right] + \\ + \left[\left(\frac{N}{\hat{N}_g}\right)^2 \left(\frac{N_h}{N}\right)^2 \left(\frac{N_h - n_h}{N_h n_h}\right)\frac{n_{h1}}{n_h}\left(1 - \frac{n_{h1}}{n_h}\right)\left(1 - \hat{P}_{g,c}\right)^2\right] \tag{7.8}$$

For revised maps, the accuracies of interest can be calculated through domain estimation, as presented by Cochran (1977, Chapter 5) and Särndal *et al.* (1992, Chapter 10). Suppose that:

N_{hd} – number of pixels in the basis map category *h* that correspond to revised map domain *d*;

N_d – number of pixels in revised map domain *d*;

N_{d1} – number of pixels in the revised map domain *d* that correspond to same reference category;

n_{hd} – number of pixels in the sample collected in basis map category *h* that intersect revised map domain *d*;

n_{hd1} – number of pixels in the sample collected in basis map category *h* that correspond to domain *d* and reference category *d*.

Overall accuracy estimation for the revised map is

$$\hat{P}_{c,n} = \frac{1}{N}\sum_{h=1}^{H} N_h \sum_{d=1}^{D} \frac{n_{hd1}}{n_{hd}} \tag{7.9}$$

and the respective estimated variance

$$\hat{V}(\hat{P}_{c,n}) = \sum_{h=1}^{H}\left(\frac{N_h}{N}\right)^2 \frac{N_h - n_h}{N_h n_h}\left[\sum_{h=1}^{H}\frac{n_{hd1}}{n_{hd}}\left(1 - \sum_{h=1}^{H}\frac{n_{hd1}}{n_{hd}}\right)\right] \tag{7.10}$$

User's accuracy estimation for each domain in the revised map is

$$\hat{P}_{d,c} = \frac{1}{N_d}\sum_{h=1}^{H}\frac{N_{hd}}{n_{hd}} n_{hd1} \tag{7.11}$$

and the respective estimated variance

$$\hat{V}(\hat{P}_{d,c}) = \sum_{h=1}^{H} \left(\frac{N_{hd}}{N_d} \right)^2 \frac{N_h - n_h}{N_{hd} n_h} \left[\frac{n_{hd1}}{n_{hd}} \left(1 - \frac{n_{hd1}}{n_{hd}} \right) \right] \qquad (7.12)$$

The estimation of producer's accuracy for the revised map, $\hat{P}_{g,cn}$, and respective estimated MSE, can be simply derived by employing formulas (7.7) and (7.8) with the correct substitutions.

We can now write the approximate confidence intervals for each generic estimate $\hat{P}$ as $\hat{P} \pm Z_{\alpha/2} \left[V(\hat{P}) \right]^{1/2}$, where $Z_{\alpha/2}$ is the standard normal deviate for the desired confidence level 1-α.

7.4. Acknowledgements

This study was carried out in the framework of the project "LANDEO - User driven land cover characterisation for multi-scale environmental monitoring using multi-sensor earth observation data (PDCTE/MGS/49969/2003)" funded by "Programa Dinamizador das Ciências e Tecnologias para o Espaço" from "Fundação para a Ciência e Tecnologia", and from the "Announcement of Opportunity for the Utilisation of ERS and ENVISAT Data" from European Space Agency (ESA). Research by Hugo Carrão was founded by the "Fundação para a Ciência e Tecnologia" (SFRH/BD/18447/2004).

7.5. References

D. H. Card, "Using known map category marginal frequencies to improve estimates of thematic map accuracy", Photogrammetric Engineering and Remote Sensing, Vol. 48, pp. 431–439, 1982.

W. G. Cochran, "Sampling Techniques", 3rd edition, New York: Wiley, 1997.

R. G. Congalton and K. Green, "Assessing the Accuracy of Remotely Sensed Data: Principles and Practices", Boca Raton, FL: CRC/Lewis Press, 1999.

S. Dicks and T. Lo, "Evaluation of thematic map accuracy in a land-use and land-cover mapping program", Photogrammetric Engineering and Remote Sensing, Vol. 56, pp. 1247-1252, 1990.

M. Ginevan, "Testing land-use map accuracy: another look", Photogrammetric Engineering and Remote Sensing, Vol. 45, pp. 1371-1377, 1979.

A. Hay, “Sampling designs to test land-use map accuracy”, Photogrammetric Engineering and Remote Sensing, Vol. 45, pp. 529-533, 1979.

C. E. Särndal, B. Swensson and J. Wretman, “Model Assisted Survey Sampling”, New York: Springer-Verlag, 1992.

S. V. Stehman, “Estimating the kappa coefficient and its variance under stratified random sampling”, Photogrammetric Engineering and Remote Sensing, Vol. 62, pp. 401–407, 1996.

S. V. Stehman, “Basic probability sampling designs for thematic map accuracy assessment”, International Journal of Remote Sensing, Vol. 20, pp. 2423–2441, 1999.

S. V. Stehman, “Statistical rigour and practical utility in thematic map accuracy”, Photogrammetric Engineering and Remote Sensing, Vol. 67, pp. 727–734, 2001.

S. V. Stehman and R. L. Czaplewski, “Design and analysis for thematic map accuracy assessment: fundamental principles”, Remote Sensing of Environment, Vol. 64, pp. 331–344, 1998.

S. V. Stehman and R. L. Czaplewski, “Introduction to special issue on map accuracy”, Environmental and Ecological Statistics, Vol. 10, pp. 301–308, 2003.

J. D. Wickham, S. V. Stehman, J. H. Smith and L. Yang, “Thematic accuracy of the 1992 National Land-Cover Data for the western United States”, Remote Sensing of Environment, Vol. 91, pp. 452–468, 2004.

M. Wulder, S. Franklin, J. White, J. Linkes and S. Magnussen, “An accuracy assessment framework for large-area land cover classification products derived from medium-resolution satellite data”, International Journal of Remote Sensing, Vol. 27, pp. 663-683, 2006.

Z. Zhu, L. Yang, S. V. Stehman and R. L. Czaplewski, “Accuracy assessment from the US Geological Survey Regional Land Cover Mapping Program: New York and New Jersey Region”, Photogrammetric Engineering and Remote Sensing, Vol. 66, pp. 425–1435, 2000.

8 A REFERENCE SAMPLE DATABASE FOR THE ACCURACY ASSESSMENT OF MEDIUM SPATIAL RESOLUTION LAND COVER PRODUCTS IN PORTUGAL

Carrão, H., Araújo, A., Cerdeira, C., Sarmento, P., Capão, L. and Caetano, M. (2007), A reference sample database for the accuracy assessment of medium spatial resolution land cover products in Portugal. In *Proceedings of the IEEE International Conference on Geoscience and Remote Sensing Symposium* (IGARSS´2007), 20-29 July 2007, Barcelona, Spain.

Abstract

This paper introduces a reference sample database produced by the Remote Sensing Unit of the Portuguese Geographic Institute for the accuracy assessment of medium scale land cover products in Portugal. The goal is to provide the worldwide remote sensing community with sufficient data for the accurate estimation of overall and per class proportions of correctly classified area in regional and global land cover products at this part of the globe. This is a massive database that encloses various descriptive attributes for each sample observation, namely primary and alternate reference land cover labels, nominally scored interpretation ratings and location confidence ratings. We present in detail the attached land cover nomenclature and the database design, i.e. the sampling design used to collect sample observations, as well as the process used to identify the most pertinent reference land cover label for each observation. In addition, we briefly describe some statistics about attribute information that was recorded for each observation and that can be used as auxiliary information for the accuracy assessment of land cover products.

Keywords: reference sample; accuracy assessment; regional and global land cover products; MERIS; MODIS; Portugal.

8.1. Introduction

The increasing availability of remotely sensed data, frequently acquired by medium spatial resolution sensors, such as the MEdium Resolution Imaging Spectrometer (MERIS) and the Moderate Resolution Imaging Spectroradiometer (MODIS), yields the increasing production of land cover maps at regional and global scales. These products serve as primary input data for several land surface models and are extremely valuable for the study of land cover

change resulting from deforestation and environmental degradation that violate international treaties, such as the Kyoto Protocol [1].

Even though low and medium scale land cover production is increasing, most having great significance in a wide area of applications, standard operational protocols for their accuracy assessment do not exist [2]. Accuracy assessment estimates general and specific qualities of the derived land cover products so that users may evaluate their utility for specific applications [3, 4]. Users should not take a map at face value without some associated estimate of error [5]. Thus, accuracy assessment should be an integral component of any mapping project based on remote sensing data. However, accuracy assessment is very often neglected due to several constrains, namely: 1) unavailability of updated and date consistent high resolution data for reference land cover appraisement; and 2) accuracy assessment is timely and economically expensive.

In this paper we present a reference sample database developed by the Remote Sensing Unit (RSU) of the Portuguese Geographic Institute (IGP) for the accuracy assessment of medium spatial resolution land cover products in Portugal. The database was initially conceived to support an ongoing research work that aims at developing a systematic classification methodology for the regular production of user driven land cover products from medium spatial resolution satellite imagery. However, this effort was lengthened in order to provide reliable and up to date reference sample data to be used in the framework of other national and even regional/global land cover programs. The goal is to focus interest in the accuracy assessment of land cover products, thus reinforcing the importance in the dissemination of reliable land cover data to end users. The database was produced for the reference year of 2005 and each stored observation represents the prevailing land cover occupying a square area of 300 m at Earth's surface at that time. We are currently adapting this reference database for 500 m pixels area. The effective application of this database has already been experienced in the comparison of land cover products derived from MERIS and MODIS single date images [6] and on the validation of a land cover map derived from the classification of multi-temporal MERIS imagery [7]. Next we introduce the land cover nomenclature, the database design, and provide a brief analysis of database features.

8.2. Land Cover Nomenclature

Database observations match up to a certain geographical spot within the Portuguese territory and are organized according to the land cover that best portray that area. There are 23 classes composing the adopted nomenclature, so far entitled LANDEO, which were used to describe land cover within each sample observation. The cover types were planned by the

RSU to be used in the production of land cover maps at national scale, and were defined through the Land Cover Classification System (LCCS) from Food and Agriculture Organization (FAO) [8] (Table 8.1). The rationale behind the development of the nomenclature was three-fold: (1) a nomenclature that is well adapted to the type of landscapes existent in regions with characteristics similar to the Portuguese mainland, (2) a nomenclature that is compatible with established ones (e.g. CORINE Land cover) in order to turn possible the comparison between our maps and others using different nomenclatures, and (3) a nomenclature that matches the spatial resolution of used satellite imagery.

Table 8.1. Land cover classes and respective number of collected sample observations. See Table 8.2 for columns titles.

Land Cover Class	**L1**	**L2**	**MC1**	**MC2**
11. Continuous Artificial Areas	25	26	4	3
12. Discontinuous Artificial Areas	43	40	64	62
21. Rainfed Herbaceous Crops	113	79	88	109
22. Irrigated Herbaceous Crops	38	32	66	45
23. Rice Crops	29	25	12	0
241. Permanent Evergreen Crops (Trees or Shrubs)	20	23	46	59
242. Permanent Deciduous Crops (Trees or Shrubs)	15	18	33	27
25. Mosaic Cultivated and Managed Lands	76	76	-	-
311. Broadleaved Closed Trees	85	67	31	16
312. Broadleaved Open Trees	76	85	35	55
321. Needleleaved Closed Trees	99	80	54	24
322. Needleleaved Open Trees	27	34	6	8
331. Mixed Closed Trees	46	49	58	30
332. Mixed Open Trees	30	42	18	18
34. Closed to Open Shrubland	203	211	110	126
35. Closed to Open Herbaceous Vegetation	62	107	36	56
36. Sparse Vegetation	31	34	6	14
38. Recently Burnt	62	38	26	25
37. Mosaic Trees/Shrubs/Herbaceous	188	188	-	-
41. Mosaic Cultivated and Managed Lands/Natural and Semi-natural areas/Artificial Areas	459	459	-	-
5. Permanent Wetlands	13	13	4	5
6. Barren	8	22	8	13
7. Water Bodies	152	152	18	28

8.3. Database Design

The process of doing the reference database for the accuracy assessment of medium spatial resolution land cover products in Portugal involved a detailed review of the state of the art. This phase guide us in the definition of a number of fundamental steps required to accommodate the necessary information for future applications of the database. As a result,

the database production process can be described as consisting of two main parts: 1) the survey sampling, namely the definition of the reference sampling observation and the selection of the most adequate sampling design for geographical random sample collection; and 2) the so-called response design [3], which consists on the evaluation of the variable of interest at each sampling observation, i.e. reference land cover labels identification by visual analysis of high spatial resolution aerial imagery. Next we detail each of these components, providing a synopsis of the main aspects taken into account during database construction.

8.3.1. Reference sampling observation

The reference sample observation serves as the basic element of comparison between the land cover map classification and the reference, or "true" classification. Sample observations can be a point or an area and their various forms and relative advantages and disadvantages for land cover maps accuracy assessment are examined in detail by [3]. In the context of large-area land cover cartography assessment, derived from pixel-based classification of medium spatial resolution satellite imagery, the sample observation is usually the minimum mapping unit of the map (i.e. pixel), which dimension is identical to the spatial resolution of the source imagery [9, 10]. In accordance, each sampling observation in the database was labeled as the most distinguishable land cover class at a specific geographical spot representing a nominal area of 300-by-300 m at Earth's surface. The selection of a reference sample observation with this specific square area was established to be in accordance with the original purpose of the database, i.e. to evaluate the accuracy of national land cover products derived from 300m spatial resolution MERIS Level 2 Full Resolution imagery.

8.3.2. Sampling design

The first step in this phase was to define the appropriate technique to geographically select the reference sample observations that will constitute the reference sample. The set of rules that must be used in this process is entitled sampling design [11]. In a random sampling design the inclusion probabilities are known for all elements in the sample and are non-zero for all elements in the population [4]. A statistically rigorous sampling random design (e.g. simple random, stratified random, cluster, and systematic sampling) contributes to a scientifically defensible accuracy assessment of land cover products, and such designs must be used because of their objectivity [3]. If this is not the case, overall and per class accuracies estimates become entirely dependent on assumptions and, as such, will be very difficult to defend in a confrontational setting [4]. Basic sampling designs, such as simple random sampling, are frequently appropriate for land cover maps accuracy assessment [12]. Often, however, it is impractical to follow such sampling procedures. This is because there is

a need to ensure specified sample sizes per land cover class to achieve more precise accuracy estimates for rare strata, and sample does not ensure that all classes are adequately represented. Stratified random sampling must be applied when there is a need to ensure a minimum sample size in each stratum to derive precise accuracy estimates for all land cover classes presented in the map [11]. Moreover, stratifying by mapped land cover classes may be used to achieve more precise estimates than would be obtained from simple random or systematic sampling without stratification [4]. The presented complexity of this process is that if design includes an unequal probability sampling, e.g. stratified sampling with equal or optimal allocation, then the inclusion probabilities of observations determined by the design must be known and incorporated in the estimation. If the inclusion probabilities are known and its presence is recognized, then the consistency criterion of statistical rigor is readily satisfied in practice by incorporating them in the accuracy estimation [3, 11, 13].

To produce a reference database large enough to ensure precise per class accuracies estimation for all cover types presented in LANDEO nomenclature, it was necessary to perform a stratified random sampling selection of observations. The difficulty inherent to this approach was the a priori inexistence of a land cover map including the same classes to be used as strata. Thus, we decided to automatically produce a primary map based on MERIS Level 2 Full Resolution imagery data from August 2005 to serve as strata for sample observations allocation. This map was produced using a set of approximately thirty sample observations per class that served as input training data of a Maximum Likelihood (ML) classifier. Overall map accuracy was assessed using an independent sample of approximately the same size of the training set and estimated as about 63%.

The next step was to define the appropriate sample size to be collected per mapped land cover type. In general, the larger the sample size the greater the confidence one can have in assessments based on that sample. Several authors have suggested methods and guidelines to estimate the appropriate sample size. The goal is to provide accuracy estimations preferably with small variance estimates at some high level of confidence. Along years, main authors have used an equation based on the binomial approximation to the normal distribution to estimate the required sample size for the accuracy assessment of land cover maps (e.g. [14]). Reference [13] states that this approach is statistically sound for estimating the sample size needed to estimate the overall accuracy of a classification or the accuracy of a single land cover category.

An absolute precision of 10% in per class accuracy estimation is sufficient to prosecute our requisites. This implies a collection of 100 sample observations per class to provide that

precision at 95% level confidence, even if per class accuracy is 0.5 [15]. Thus, we guaranteed that 100 sample observations were collected per mapped class. Note that for the accuracy assessment of upcoming maps with different cover types we can still use this data to estimate overall and per class accuracies. However, absolute precision of estimates may be larger than 10%, according to the spatial distribution of those cover types, from now on entitled domains. For a detailed description on how to estimate overall and per class accuracies based on domain estimation, please read [16].

8.3.3. Reference imagery data

The identification of the cover type most typical in each sample observation requires the visual analysis of high spatial resolution remote sensed imagery or even field check. The former is timely and economically expensive, so we decided to use high spatial resolution aerial imagery to identify the most adequate land cover type at each sample observation. Reference land cover labels were derived by visual analysis of orthorectified aerial images acquired during the years of 2004, 2005 and 2006 and covering the whole Portuguese territory; these images have the following characteristics: four spectral bands in the blue, green, red, and infrared wavelengths; and 50 cm spatial resolution. The bounds of sample observations, randomly selected for each cover type, were identified using a 300-by-300 m fishnet covering the entire Portuguese territory.

8.3.4. Image interpretation

This phase may be considered as the most important and frail of the reference sample database production. At this stage, four image interpreters were escorted in the identification of the most reliable land cover type characterizing each sample observation. Reference [12] suggests that an uncertainty in the confidence of single class labeling using the available reference data can significantly influence the apparent accuracy of a classification. Thus, to avoid confounding classification and location errors, primary and alternate reference land cover labels must be identified and characterized by visual interpretation of aerial images for each reference sample observation. This process was done in a blind format in that the interpreters did not have knowledge of the primary map classification for each observation. After image interpretation training, performed to reduce individual subjectivity in reference label assignment, image interpreters collected primary and alternate reference land cover labels for sample observations and surrounding pixels. This process was executed repeatedly for each observation by all image interpreters to reduce uncertainty whenever the label identification was not clear. When this happened, image interpreters altogether examined critically the situation and jointly decided for the most adequate label. Additional auxiliary

information, such as intra-annual time series of Normalized Difference Vegetation Index for each sample observation, was used to complement the fragile visual analysis of single date aerial imagery. Nominally scored interpretation and location confidence ratings (ICR and LCR, respectively) were also specified for each observation, as suggested by [10]. These ratings allow the use of a scale of error in the accuracy assessments rather than simple and perhaps overly severe interpretations of ground data [10]. In Table 8.2 we present a description of the label and auxiliary information collected for each sample observation recorded in the reference database.

8.4. Database Analysis

In this section we briefly present some statistics describing the reference label of sample observations integrated in the database and of their respective attributes. Recurring to Table 8.1, we perceive that the number of observations collected per reference class is quite different from the 100 established per map class. This is due to some classification errors in the primary map that was used as strata, but it will not affect the effective utilization of the reference database in future applications. The repercussion is that absolute precision of accuracy estimations for other maps may be larger than the foreseen 10%, as stated by [16]. Independently, these outcomes reveal the importance of the stratified sampling design for sample selection. Some classes, such as Continuous Artificial Areas and Barren, are not too abundant in the Portuguese landscape, and would not be prevalent in the final sample if a simple random sampling design has been used for its selection.

Table 8.2. Attribute Information Collected for Interpreting Agreement Between Map and Reference Data. Adapted from [12].

Information from Reference Source
(1) Primary reference label (L1): land-cover label thought to be most correct by image interpreter.
(2) Alternate reference label (L2): land-cover label that might also be considered correct given information in the aerial image. An alternate reference label must not be provided if the image interpreter's primary reference label is believed unambiguously evident.
(3) Most abundant cover type within a mosaic observation (MC1) and second most abundant cover type in the same mosaic observation (MC2).
(4) Other information on land cover, such as: clear cuts; mineral extraction sites; construction sites; and salines.
(5) Image interpreter confidence rating (ICR): nominal ranking of image interpreter's certainty in identifying the primary reference land-cover label. Rank values range from 1 (not confident) to 3 (very confident).
(6) Location confidence rating (LCR): nominal score for location of sample pixel: 1 = on the edge between two land-cover classes; 2 = homogeneous area of land cover; and 3 = heterogeneous area of land cover.
(7) Date: day, month, and year of aerial image acquisition.
(8) Notes: entries for any other factors that may affect image interpretation (e.g., temporal change).

We now proceed with the analysis of ICR and LCR (Table 8.3). Statistics show that reference label interpretation of sample observations was very often considered as “very confident”. This is due to the fact that all sample observations were successively interpreted by the four image interpreters, thus reducing the uncertainty in reference labeling. Location confidence ratings do not present any predisposition, reflecting the fragmentation of the Portuguese landscape, ranging from homogeneous to expressively heterogeneous land cover areas at a 300-by-300 m spatial resolution. However, this may anticipate a truly difficult task, i.e. the automatic characterization of the Portuguese land cover with medium spatial resolution satellite imagery.

Table 8.3. ICR and LCR statistics for reference database

Code	ICR	LCR
1	396	632
2	164	707
3	1340	561

8.5. Conclusion

A reference sample database for the accuracy assessment of medium scale land cover products was produced for the conterminous Portuguese territory. The development phase is concluded, but the database will be continuously updated – depending on the availability of reference data – to be timely effective. This database is presented in the website of RSU (www.igeo.pt/gdr) and can be provided by request.

8.6. Acknowledgments

This study was carried out in the framework of the project "LANDEO - User driven land cover characterisation for multi-scale environmental monitoring using multi-sensor earth observation data (PDCTE/MGS/49969/2003)" funded by "Programa Dinamizador das Ciências e Tecnologias para o Espaço" from "Fundação para a Ciência e Tecnologia (FCT)", and from the "Announcement of Opportunity for the Utilisation of ERS and ENVISAT Data" from European Space Agency (ESA). Research by Hugo Carrão (SFRH/BD/18447/2004) and Luis Capão were founded by "FCT".

8.7. References

[1] W. Michael, S. Franklin, J. White, J. Link, and S. Magnussen, “An accuracy assessment framework for large-area land cover classification products derived from

medium-resolution satellite data", International Journal of Remote Sensing, vol. 27, pp. 663–683, February 2006

[2] C. Justice, A. Belward, J. Morisette, P. Lewis, J. Privette, and F. Baret, "Developments in the validation of satellite sensors products for the study of the land surface", International Journal of Remote Sensing, vol. 21, pp. 3383–3390, 2000.

[3] S. Stehman, and R. Czaplewski, "Design and analysis for thematic map accuracy assessment: fundamental principles", Remote Sensing of Environment, vol.64, pp. 331–344, 1998.

[4] S. Stehman, "Statistical rigour and practical utility in thematic map accuracy", Photogrammetric Engineering and Remote Sensing, vol. 67, 727–734, 2001.

[5] D. Card, "Using known map category marginal frequencies to improve estimates of thematic map accuracy", Photogrammetric Engineering and Remote Sensing, vol. 48, pp. 431–439, 1982.

[6] H. Carrão, P. Sarmento, A. Araújo, and M. Caetano, "Retrieving land cover information from MERIS and MODIS data: a comparative study for landscape characterization in Portugal", Proceedings of the 2007 International Geoscience and Remote Sensing Symposium (IGARSS´2007), 2007.

[7] L. Capão, H. Carrão, A. Araújo, and M. Caetano, "An approach for land cover mapping with multi-temporal meris imagery", Proceedings of the 2007 International Geoscience and Remote Sensing Symposium (IGARSS´2007), 2007.

[8] Di Gregorio, L. Jansen, "Land cover classification system LCCS: Classification concepts and user manual", FAO, Rome, 2000.

[9] Z. Zhu, L. Yang, S. Stehman, and R. Czaplewski, "Accuracy assessment from the US Geological Survey Regional Land Cover Mapping Program: New York and New Jersey Region", Photogrammetric Engineering and Remote Sensing, vol. 66, pp. 1425–1435, 2000.

[10] J. Wickham, S. Stehman, J. Smith, and L. Yang, "Thematic accuracy of the 1992 National Land-Cover Data for the western United States", Remote Sensing of Environment, vol. 91, pp. 452–468, 2004.

[11] S. Stehman, "Basic probability sampling designs for thematic map accuracy assessment", International Journal of Remote Sensing, vol. 20, pp. 2423–2441, 1999.

[12] G. Foody, “Status of land cover classification accuracy assessment”, Remote Sensing of Environment, vol. 80, pp. 185–201, 2002.

[13] R. Congalton, and K. Green, “Assessing the Accuracy of Remotely Sensed Data: Principles and Practices”, Boca Raton, FL: CRC/Lewis Press, 1999.

[14] Hay, “Sampling designs to test land-use map accuracy”, Photogrammetric Engineering and Remote Sensing, vol. 45, pp. 529–533, 1979.

[15] W. Cochran, “Sampling Techniques”, 3rd edition, New York: Wiley, 1997.

[16] H. Carrão, M. Caetano, and P. Coelho, “Sample Design and Analysis for Thematic Map Accuracy Assessment: An Approach Based on Domain Estimation for the Validation of Land Cover Products”, Proceedings of the 32nd International Symposium on Remote Sensing of Environment, 2007.

9 MULTITEMPORAL MERIS IMAGES FOR LAND COVER MAPPING AT NATIONAL SCALE: THE CASE STUDY OF PORTUGAL

Carrão, H., Araújo, A., Gonçalves, P. and Caetano, M. (2008). Multitemporal MERIS images for land cover mapping at national scale: the case study of Portugal. *International Journal of Remote Sensing*, to be published.

Abstract

Medium spatial resolution satellite images already proved successful in the automatic production of global land cover maps. However, their operational use for land cover mapping at national scales was not yet well succeeded. We claim that reasons are not data source dependent, but due to adopted land cover nomenclatures properties, regional landscapes specificities and used methodological approaches. The goal of this paper is to evaluate the suitability for national applications of land cover maps derived from automatic classification of medium spatial resolution satellite images. To tackle this issue, we produced a land cover map of Continental Portugal from multitemporal MEdium Resolution Imaging Spectrometer (MERIS) Full Resolution satellite images of 2005 and evaluated its accuracy. For the accuracy assessment of the final map we computed unbiased estimates of overall, user's and producer's accuracies using an independent testing sample collected through a stratified random sampling design. Overall accuracy of final map was of 80% with an absolute precision of 2% at the 95% confidence level. High independent accuracy assessment results demonstrate that medium spatial resolution satellite images can be used on an operational basis for annual production of land cover maps suitable for national applications.

Keywords: Land cover classification; multitemporal MERIS images; Linear Discriminant Classifier; Portugal.

9.1. Introduction

Automatic production of land cover maps for global applications is in constant expansion. The reasons are clear: medium spatial resolution sensors, such as the Moderate Resolution Imaging Spectroradiometer (MODIS) and the MEdium Resolution Imaging Spectrometer (MERIS), exhibit high spectral and temporal resolutions, wide geographical coverage and improved atmospheric correction. Indeed, the large number of spectral bands turns it

possible to identify more detailed land cover classes with higher accuracy than would be possible with data from elder sensors with similar temporal and spatial resolutions, such as the Advanced Very High Resolution Radiometer (AVHRR) (Borak and Strahler 1999). Besides, they also outperform sensors collecting spectral data at higher spatial resolutions, such as Landsat, which are not adequate to automatically and efficiently derive comparable land cover products for large areas (DeFries *et al.* 2000). Data from satellites with higher spatial resolutions have incomplete spatial coverage, infrequent temporal coverage with inevitable cloud contamination, and the associated large volumes are not practicable in an automatic land cover classification context.

MODIS data already proved successful in the production of a global land cover map, MODIS Land Cover (MOD12Q1), with principal objective to improve our understanding of the extent and of the distribution of major worldwide land cover types (Hodges 2001, Friedl *et al*. 2002). Similarly, MERIS data will be used to produce a global land cover map for year 2005/2006 in the framework of the GlobCover initiative (Arino *et al.* 2007). This land cover product of 300 m spatial resolution will update and complement other existing global land cover products, such as the Global Land Cover 2000 (GLC2000) derived from Systeme Probatoire d'Observation de la Terre (SPOT) VEGETATION daily 1-km data for the year 2000 (Bartholomé and Belward 2005).

Although these land cover maps were optimized for global applications (Giri *et al.* 2005), they are regularly updated and used as primary input data for national applications. For instance, it is the case for analysis of land cover changes that violate international treaties, like Kyoto Protocol (Wulder *et al.* 2006). However, the use of global land cover products for particular areas of interest may not meet users' requirements. According to Giri *et al.* (2005), global land cover products present high overall mismatch (41%) with detailed land cover classes and moreover agreement is highly variable among regions. Such divergence can be explained by differences in the definition of land cover classes, input data sources, classification methodologies, and spatial details used in image classification (Giri *et al.* 2005). Moreover, it is reckoned that these global products present smaller overall accuracies for certain particular areas of interest than depicted by producers. For example, MOD12Q1 producers claim overall accuracies between 75-80% (Giri *et al.* 2005), whereas the validation of this product for Portugal does not exceed 57% accuracy (Cerdeira *et al.* 2006). In addition, comparison of MOD12Q1 with Co-ordination of Information on the Environment (CORINE) Land Cover map for 2000 (CLC2000, EEA 2002), within the same region, showed an only 50% agreement between the two products (Caetano and Araújo 2006).

To meet users' needs for regional and national applications, several research studies focused on the automatic production of land cover maps from medium spatial resolution satellite images. MODIS data was used in several regional experimental applications: production of a landform map of North Africa (Ballantine *et al.* 2005), regional land cover mapping to monitor biological conservation in the Great Yellowstone Ecosystem (Wessels *et al*. 2004), and land cover classification over the Yellow River basin, China (Matsuoka *et al*. 2004), to cite but a few. Similarly and most recently, MERIS data was also used for: land use-land cover mapping in Spain (García-Gigorro *et al.* 2007), land cover mapping in the Netherlands (Clevers *et al.* 2007), and land cover mapping in Wisconsin, USA (Dash *et al.* 2007). Indeed, automatic classification is required because operational land cover programs at national scales, that resort to visual interpretation and manual classification of high spatial resolution remote sensing data, such as CORINE Land Cover (EEA 2002), are time consuming and economically expensive.

The efforts made towards development and description of automatic techniques for land cover classification are promising, but the accuracy assessment of derived maps is either insufficient or incorrectly reported. Land cover products are frequently assessed using "unseen training sites", e.g. Dash *et al.* (2007). In this situation, map accuracy measures become entirely dependent on assumptions and, as such, do not provide users with unbiased estimates of map overall and per class accuracies (Stehman 2001). Similarly, rather than an accuracy assessment exercise, authors often compare their land cover maps with other existing land cover products in order to identify differences and/or obvious errors, e.g. García-Gigorro *et al.* (2007) and Clevers *et al.* (2007). This approach, referred to as a "confidence-building approach" (Strahler 2002), does not provide users with information about the estimated accuracy of the final map. Hence, frequently used accuracy assessment designs are neither sufficient to properly evaluate the true accuracy of derived land cover maps nor statistically sound.

In this paper we evaluate the suitability for national applications of land cover maps derived from automatic classification of medium spatial resolution satellite images. Although effective accuracy assessments of global land cover maps for particular areas of interest revealed inadequate for national applications, we believe that data source is not primarily responsible. Indeed, we claim that origins are ground in nomenclature properties, regional specificities, used methodological approaches, and more precisely to inadequate training data sets. To support our claim, we derived a land cover map for Continental Portugal from automatic classification of multitemporal MERIS images for the year 2005, and evaluated its

accuracy through a random sampling design. Our methodological approach decomposes in three steps:

- Definition of a land cover nomenclature that fits well the landscape characteristics of our study area and that can be easily paired with other nomenclatures;
- Independent classification techniques for intra-annual stationary (e.g. artificial areas) and for transient (e.g. burnt areas and clear cuts) land cover classes;
- Adequate estimation of final map overall and per class accuracies.

To complement our map evaluation, we compared it with a reliable land cover map derived for the Portuguese territory from visual analysis of high resolution satellite data (CORINE Land Cover map for 2000). This study is the comprehensive outcome of previous experimental results attained by the Remote Sensing Unit (RSU) of Portuguese Geographic Institute (IGP):

- Definition of a land cover nomenclature well adapted to Portuguese territory (Carrão *et al.* 2007a).
- Evaluation of Linear Discriminant Classifier (LDC) for land cover mapping with adequate training data (Carrão *et al.* 2006a).
- Classification gain by using multitemporal imagery data (Carrão *et al.* 2007b, 2007c).
- Use of vegetation index differentiating techniques for accurate mapping of areas with intra-annual vegetation decreases (Armas and Caetano 2005).
- Application of a statistical random sampling design for the accuracy assessment of medium spatial resolution land cover maps (Carrão *et al.* 2007d).

9.2. Land cover nomenclature

The Remote Sensing Unit (RSU) of Portuguese Geographic Institute (IGP) has recently recommended a land cover nomenclature (Carrão *et al*. 2007a) to be used in regular production of user driven land cover products for multi-scale environmental monitoring in Portugal. The adopted hierarchical nomenclature, entitled LANDEO, encompasses 16 comprehensive land cover classes at the most disaggregated level (Table 9.1). These classes were defined using the Land Cover Classification System (LCCS) from Food and Agriculture Organization (FAO) (Di Gregorio and Jansen 2000). LANDEO nomenclature includes land cover types that are considered as stationary along the year, e.g. artificial areas, and others that undergo transient changes, such as burnt areas and clear cuts. In this nomenclature, the number and the type of classes are comparable to that of the Global Land Cover (Bartholomé and Belward 2005) and the International Geosphere-Biosphere

Programme (Belward *et al.* 1999) nomenclatures. The intentions behind LANDEO development were three-fold: (1) a nomenclature that is well adapted to the characteristics of landscapes existent in regions similar to Portuguese territory; (2) a compatible nomenclature with established ones to permit comparison with other mapping products; and (3) a nomenclature that can be used with different spatial resolution satellite images (Carrão *et al.* 2006a).

Table 9.1. LANDEO land cover nomenclature at the most disaggregated level.

Label	Land cover class	Class description
5	Wetlands	The land cover consists of a permanent mixture of water and vegetation. The vegetation can be present in salt, brackish, or fresh water.
6	Bare to sparsely vegetated areas	The land cover consists of natural areas with less than 15% of vegetation cover during all time of year. It includes areas like bare rock and sands.
7	Water	The land cover consists of natural/artificial water bodies. Can be either fresh or salt-water bodies.
11	Continuous artificial areas	The land cover consists of artificial areas (e.g. buildings, roads). At least 80% of the total surface must be impermeable.
12	Discontinuous artificial areas	The land cover consists of artificial areas (e.g. buildings, roads). Between 30-80% of the total surface must be impermeable.
21	Non-irrigated herbaceous crops	The land cover consists of rainfed herbaceous crops. These crops are annually harvested and followed by a bare soil period.
22	Irrigated herbaceous crops	The land cover consists of irrigated herbaceous crops. These crops are annually harvested and followed by a bare soil period.
23	Rice crops	The land cover consists of rice crops. These crops are annually harvested and followed by a bare soil period.
242	Vineyards	The land cover consists of permanent deciduous crops, namely vineyards.
34	Shrubland	The land cover consists of woody vegetation (shrubs) with more than 15% cover and with less than 5 m height. Tree cover is less than 40%.
35	Grassland	The land cover consists of natural and cultivated herbaceous vegetation for livestock feeding with more than 15% cover. Tree and shrub cover is less than 40%.
38	Burnt areas and clear cuts	The main layer consists of closed to open trees or shrubs affected by forest fires or clear cuts.
311	Broadleaf forest	The land cover consists of broadleaf trees with at least 5 m height and with a crown cover of more than 40%.
312	Agro-forestry areas	The land cover consists of broadleaf trees with at least 5 m height and with a crown cover between 15-40% and with understory agricultural systems.
321	Needleleaved forest	The land cover consists of needle leaf trees with at least 5 m height and with a crown cover of more than 40%.
331	Mixed forest	The land cover consists of a mixture or mosaic of broadleaf and needle leaf trees with at least 5 m height and with a crown cover of more than 40%. None of the tree types exceeds 80% of the mixture or mosaic.

Regarding this last intention, let us mention though, that our nomenclature does not include mixed classes. However, at coarse imagery spatial resolutions land cover homogeneity is hardly guaranteed within a pixel area, hence most pixels are likely to contain features from two or more distinct homogeneous classes (Carrão *et al.* 2007b). Thus, we decided, in

accordance with Cihlar *et al.* (1996) and Hansen *et al.* (2000), to assign mixed areas to the dominant land cover type within a pixel and not to classify them as mixed.

9.3. Study area

We focused our experiences on the Portuguese mainland territory. The country has an area of approximately 89 000 km^2 and ground altitudes ranging from 0 to 2000 m above sea level. According to Caetano *et al.* (2005), in 2000 the agricultural areas occupied 37% of the country, closely followed by semi-natural areas (33%) and forest (27%), while artificial surfaces comprise only 3% of the total land cover. Portugal is in a transition region featuring diverse landscapes representing both Mediterranean and Atlantic climate environments.

9.4. Dataset

The data used to perform this study are divided in: (1) training and testing data: MERIS images collected for automatic mapping of land cover types; and (2) ancillary data: existing land cover cartographies and aerial images used as complementary information for reference land cover type's identification.

9.4.1. Source data

MERIS sensor is on board of Environmental Satellite (ENVISAT), an advanced polar-orbiting Earth Observation (EO) satellite launched by the European Space Agency (ESA) in 2002. This sensor allows a global coverage of the Earth in three days, and provides the most radiometrically accurate data on Earth's surface that is currently acquired from space (Curran and Steele 2005). MERIS data consist of calibrated surface reflectances in 15 groups of wavelengths (ranging from the visible (390 nm) to the near infrared (1040 nm)). In addition, they contain geophysical parameters and quality flags provided on a pixel basis, which address the quality and the validity of surface reflectances.

In this study we exploited two different sets of MERIS images with 300 m nominal spatial resolution. Both sets of images were atmospherically corrected for Rayleigh scattering, ozone, water vapour absorption, aerosol content and different illumination and viewing geometries. For stationary land cover classes' identification, we used a set of 6 bimonthly MERIS Full Resolution (FR) images composites, ranging from May 2005 to April 2006, freely available from Ionia GlobCover website (http://ionia1.esrin.esa.int/). Before compositing, an additional correction scheme was applied to these images in the framework of GlobCover project. This correction consisted of improved cloud detection, atmospheric correction, geolocation and re-mapping, as fully described in Arino *et al.* (2007). Then, for each spectral band, the average surface reflectance was calculated from all valid observations

during two consecutive month's period. Bimonthly composites were derived for 13 spectral bands (original bands 11 and 15 were removed since they address O_2 content and water vapour, respectively) and for the *Normalized Difference Vegetation Index* (NDVI). This way, each sample observation is a time series of 6 bimonthly points lying in the 14 dimensional space $[0, 1]^{13} \times [-1, 1]^{1}$ corresponding, respectively, to the surface reflectances and to the vegetation index of a 300 m-by-300 m square area.

For intra-annual transient land cover classes' identification we resort to 2 NDVI maps, from September 2004 and 2005, computed from MERIS Level 2 FR images.

9.4.2. Ancillary data

As ancillary data we used: (1) orthorectified aerial images with four spectral bands in the blue, green, red and infrared wavelengths, and with 50 cm spatial resolution – these aerial images cover the whole Portuguese territory and were acquired during years 2004, 2005 and 2006; and (2) five different existing land cover cartographies that map the entire Portuguese territory: CORINE Land Cover maps of 1990 and 2000 (CLC'90 and CLC2000), 1990 Land Cover Map (COS'90), National Forest Inventory (IFN'95) and Forest Cartography (CFE'95). CLC2000 was in addition used to complement the evaluation of our final land cover map. Hence, we present in Table 9.2 the CLC2000 map nomenclature.

Table 9.2. CLC2000 level 3 nomenclature.

111. Continuous urban fabric	311. Broad-leaved forest
112. Discontinuous urban fabric	312. Coniferous forest
121. Industrial or commercial units	313. Mixed forest
122. Road and rail network and associated land	321. Natural grassland
123. Port areas	322. Moors and heathland
124. Airports	323. Sclerophylous vegetation
131. Mineral extraction sites	324. Transitional woodland / shrub
132. Dump sites	331. Beaches, dunes, sandplains
133. Construction sites	332. Bare rock
141. Green urban areas	333. Sparsely vegetated areas
142. Sport and leisure facilities	334. Burnt areas
211. Non-irrigated arable land	335. Glaciers and perpetual snow
212. Permanently irrigated land	411. Inland marshes
213. Rice fields	412. Peatbogs
221. Vineyards	421. Salt marshes
222. Fruit trees and berry plantations	422. Salines
223. Olive groves	423. Intertidal flats
231. Pastures	511. Water courses
241. Annual crops associated with permanent crops	512. Water bodies
242. Complex cultivation patterns	521. Coastal lagoons
243. Land principally occupied by agriculture, with significant areas of natural vegetation	522. Estuaries
	523. Sea and ocean
244. Agro-forestry areas	

9.5. Methodology

The land cover map of Continental Portugal derived from MERIS FR images of 2005 was developed in five steps that we describe thereafter. Firstly, we collected two reference samples characterizing the reflectance of each land cover class within spectral wavelengths measured by MERIS sensor. These samples were used independently for training and testing purpose of our classification methodology. Afterwards, we identified and removed outlier observations from training dataset within each land cover class. Then we proceeded with the map production process in two phases. In the one hand, we produced a map with 15 intra-annual stationary land cover classes using the supervised Linear Discriminant Classifier (LDC). In the second hand, we produce a map of intra-annual transient land cover classes and overlaid it with the intra-annual stationary land cover map. The map production was concluded with a probabilistic accuracy assessment of the final product using the independent testing sample.

9.5.1. Reference samples collection

We collected two reference samples of MERIS pixels that were used separately for training and testing purpose in our classification methodology. We used different processes for training and for testing sets collection, in order to meet the specificities inherent to each of these tasks.

For each land cover class of LANDEO nomenclature, a set of training observations was deterministically selected using ancillary data as strata. Transient classes were not included in this process, since their classification was not supervised. To guarantee land cover homogeneity within training sample, classes' observations were selected preferentially in geographic spots located in the centre of 3×3 MERIS pixels of homogeneous land cover. The goal was to reduce intra-class observations dissimilarities due to possible geometric disparities between multitemporal MERIS images. Nevertheless, we endeavoured at spreading as much as possible classes' observations over the mainland territory, so as to account for possible regional classes' differences, and to avoid geographic correlation between adjacent pixels during classifier training (Hammond and Verbyla 1996).

For testing observations collection, a random sampling design was preferred. We selected one hundred observations per mapped class in order to attain a maximum absolute precision of 10% in per class accuracy estimation at the 95% confidence interval. The rationale and the statistical justification of this approach is presented in Cochran (1977, pp. 53). We selected this sampling design to ensure that all mapped classes were equally represented in the sample. Moreover, stratifying by mapped land cover classes' yields better precise overall

accuracy estimates than simple random or systematic sampling without stratification (Stehman 2001).

Following the lines of DeFries *et al.* (1998), we overlaid a co-registered 300 m fishnet corresponding to the MERIS FR data grid with the aerial images for visual interpretation of training and testing reference land cover samples. Intra-annual time series plots of NDVI were used as auxiliary information to complement the visual interpretation of land cover classes on unitemporal aerial imagery. We then instructed four aerial images interpreters to independently confirm our training observations scoring, and to identify the most sensible reference land cover class for each testing pixel. For every litigious pixel, a consensual decision was taken by all four interpreters together. However, at this spatial resolution, most (randomly selected) testing pixels are likely to contain features from two or more distinct homogeneous classes. Whenever identification of land cover types was not obvious, interpreters collected primary and secondary reference land cover classes for testing observations, similarly as in Zhu *et al.* (2000) and Wickham *et al.* (2004), to cite but a few. Primary land cover class is the one considered as the most correct by image interpreter, while secondary land cover class can also be considered correct given information in the aerial image. We considered as correct a match between map class and the primary or secondary reference land cover class of testing pixels. In Table 9.3 we present the reference sample sizes for training and testing. Note that the number of testing observations per class is not 100, because Table 9.3 relates to reference land cover classes and not to map classes used as strata for sample collection.

Table 9.3. Reference sample sizes for training and testing. Test 1 and Test 2 represent, respectively, the number of observations per land cover class collected as primary and secondary reference.

Label	**Land cover class**	**Train**	**Test 1**	**Test 2**
5	Wetlands	43	23	47
6	Bare to sparsely vegetated areas	45	64	51
7	Water	42	168	162
11	Continuous artificial areas	50	38	30
12	Discontinuous artificial areas	48	85	103
21	Non-irrigated herbaceous crops	59	174	135
22	Irrigated herbaceous crops	50	65	68
23	Rice crops	43	41	29
242	Vineyards	46	42	40
34	Shrubland	68	325	342
35	Grassland	55	338	218
38	Burnt areas and clear cuts	-	62	36
311	Broadleaf forest	48	89	115
312	Agro-forestry areas	55	154	277
321	Needleleaved forest	54	133	162
331	Mixed forest	44	99	85

9.5.2. Detecting outliers and cleaning training sample

Despite all the precision held in the collection of the training sample, it was sometimes difficult to identify the most sensible reference land cover class for some observations, even resorting to ancillary data. In this sense, it was not surprising that the preliminary training data set contained one or few unusual observations that do not seemed to belong to the land cover classes in which they were classified. Thus, we were led to develop some additional treatment for outliers' detection and removal from training sample. Johnson and Wichern (1998) state that based on a single characteristic, unusual observations data are those that are either very large or very small compared to the others. If the number of characteristics is smaller than three, outliers can be roughly detected by visual analysis of scatter plots. In the case of remotely sensed data, where each observation corresponds to an image pixel and is characterized by a *p*-dimensional vector $\boldsymbol{x} = [x_1,\ldots, x_p]^T \in R^p$ of measurements on *p* spectral bands and vegetation indices, outliers identification can be more complicated. For these higher dimensional feature spaces, a more formal method for identifying outliers is based on the squared generalized distance, or Mahalanobis distance (Johnson and Wichern 1998). Assume that there are *M* land cover classes in the adopted nomenclature, labeled C_1 to C_M, organized as a set of labels $\Omega = \{C_1,\ldots,C_M\}$, and that each sample observation belongs to one and only one land cover class. Let n_g be the number of observations in our training data set **Z** from reference class C_g, $g = 1,\ldots, M$, and $l(\mathbf{z}_j)\in \Omega$ be the class label of observation $\mathbf{z}_j\in \mathbf{Z}$. The square generalized distance, d_{jg}^2, is computed as follows:

$$d_{jg}^2 = (\mathrm{z}_j - \hat{\mu}_g)^T \hat{\Sigma}_g^{-1}(\mathrm{z}_j - \hat{\mu}_g), \; j = 1, \ldots, n_g, \text{ and } g = 1, \ldots, M. \tag{9.1}$$

The estimated C_g multispectral mean is obtained by

$$\hat{\mu}_g = \frac{1}{n_g} \sum_{l(\mathrm{z}_j)=C_g} \mathrm{z}_j \tag{9.2}$$

and the estimated covariance matrix by

$$\hat{\Sigma}_g = \frac{1}{n_g - 1} \sum_{l(\mathrm{z}_j)=C_g} (\mathrm{z}_j - \hat{\mu}_g)(\mathrm{z}_j - \hat{\mu}_g)^T. \tag{9.3}$$

Assuming that each training *p*-dimensional vector observation $\mathrm{z_j}$ comes from a land cover class C_g with multivariate normal distribution, and both the sample dimension n_g and $n_g - p$ are greater than 25 or 30, each of the squared distances d^2_{1g}, d^2_{2g}, …, d^2_{ngg} should behave like a χ^2 random variable. A prior analysis of the multivariate normality assumption can be performed for each land cover class using the method described by Johnson and Wichern (1998, pp. 199). If assumption holds, we can compare the squared generalized distances

results with the χ^2_p distribution at some 100(1-α)% confidence level. If d^2_{jg} is greater than χ^2_p (α), than we can reject the hypothesis that observation z_j belongs to the originally scored land cover class. Moreover, there exist classes that are naturally heterogeneous. We intentionally choose a rather loose selection criterion, to account for pixels that despite being borderline are characteristic of physical phenomena underlying those classes. We fix α to 0.01 and reject only observations with d^2_{jg} larger than χ^2_p (0.01).

9.5.3. Supervised classification of intra-annual stationary land cover classes

For the classification of intra-annual stationary land cover classes we resort to a Linear Discriminant Classifier (LDC). LDC is a parametric classifier, derived by approximating the class-conditional probability density functions (pdfs). Consider that spectral bands and vegetation indices values for a given pixel of unknown membership are arranged as a *p*-dimensional vector $\boldsymbol{x}$. Then, $\boldsymbol{x}$ will be mapped to the class C_g, $g = 1, \ldots, M$, that yields the highest score for the linear discriminant function h_g, defined as

$$h_g(\mathrm{x}) = \frac{1}{2}\hat{\mu}_g{}^{\mathrm{T}}\,\Sigma^{-1}\,\hat{\mu}_g + \hat{\mu}_g{}^{\mathrm{T}}\,\Sigma^{-1}\,\mathrm{x}\,. \qquad (9.4)$$

The common covariance matrix $\sum$ is estimated as the weighted average of the separately estimated class-conditional covariance matrices and reads:

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum_{g=1}^{M} n_g \hat{\boldsymbol{\Sigma}}_g \;,\; n = n_1 + \ldots + n_M. \qquad (9.5)$$

Kuncheva (2004) claims that LDC is simple to calculate from training data and is reasonably robust, i.e. the results might be surprisingly good even when the classes do not have normal distributions. Indeed, from our experiments, this classifier attain results very comparable to the ones achieved with recently developed supervised classification systems, such as Support Vector Machines (SVM) (Carrão *et al.* 2006a) or Self-Organizing Maps (SOM) (Carrão *et al.* 2006b). In addition, as LDC is computationally less expensive and involves no extra parameter to be tuned, it reveals a better choice for automatic land cover classification in an operational context (Carrão *et al.* 2006a). These results are in good agreement with the conclusions published by Wilkinson (2005) on the assessment of the degree of progress made in automatic land cover mapping between 1989 and 2003. He alludes that there is no discernible difference in the collective land cover classification accuracy values attained with parametric and non-parametric approaches.

We used the complete set of 6 bimonthly images composites as the input features for classification. It was experimentally proven in Carrão *et al.* (2007b) that combining intra-

annual multitemporal spectral information for land cover classes' discrimination in Portugal can significantly improve the best overall classification attained with single date spectral information. Moreover, in the same article it was shown that most pair-wise classes' discrimination in this geographical region benefits from the use of spectral data from at least one summer and one winter date.

9.5.4. Intra-annual transient land cover classes classification

The strategy for the classification of intra-annual transient land cover classes, corresponding to label 38 in our nomenclature (Table 9.1), follows a set of procedures defined by Armas and Caetano (2005). The methodology applied to identify areas with intra-annual vegetation decreases was the vegetation index differentiating (Lu *et al.* 2004). In this methodology, NDVI maps are produced separately for two different dates and then the vegetation index difference map is computed. The result is a new map with positive and negative values for regions where changes had occurred, and null values for areas of no change. However, due to the fact that small differences can occur because of less significant changes of vegetation, it is necessary to carefully define a difference threshold, over which relevant changes are identified (Jensen 1996). For the threshold definition, a statistical approach based on the normal distribution of the difference map is usually used. In this approach, it is considered that a pixel had considerably changed if its NDVI difference exceeds the mean value of the change map by "*m*" (to be defined) times the standard deviation (Salvador *et al.* 2000).

To identify burnt areas and clear cuts we produced separately two NDVI maps from MERIS Level 2 FR images of two consecutive years (September 2004 and September 2005) and applied the change detection method based on these two maps. As the main goal at this phase was to detect decreases in forested and natural vegetated areas, we derived an analysis mask of these land cover classes from the intra-annual stationary land cover classes' classification, and determined land cover changes only for these areas. Armas and Caetano (2005) empirically identified the best pre-factor "*m*" to detect changes in forested and natural vegetated areas in Portugal, using satellite images of medium spatial resolution (i.e. ranging from 250 to 300 m). Accordingly, we set the value of "*m*" to 1.25. Moreover, Armas and Caetano (2005) state that the best compromise between commission and omission errors in detecting vegetation decreases is attained with a minimum mapping unit of ten contiguous pixels. Thus, we have discarded all changes smaller than 90 ha from our final map. The exclusion of these areas was not penalizing because they constituted less than 10% of 2005 total burnt area in the Portuguese territory (DGRF 2005).

9.5.5. Accuracy assessment of final map

In order to evaluate the accuracy of the final map, standard measures, i.e. overall accuracy (P), user's accuracy (P_h) and producer's accuracy (P_g), were estimated from the classification of the testing sample. We followed the set of procedures presented by Card (1982) to properly estimate those accuracy measures under a stratified random sampling design. To improve the precision of our estimates, we used the whole map classification as an auxiliary variable in the estimation procedure, as described in Carrão *et al.* (2007d). Hence, land cover classes occupying larger areas in the map have a larger effect in the overall accuracy estimation. Let us fix some notations:

N – total number of pixels in the map;

N_h – total number of pixels in the map category C_h;

N_g –total number of pixels in the reference category C_g;

N_{hg} – total number of pixels in the map category C_h that intersect reference category C_g;

N_{hh} – total number of pixels in the map category C_h that intersect reference category $C_{g=h}$;

n_h – number of pixels in the test sample collected in map category C_h;

n_g – number of pixels in the test sample collected in reference category C_g;

n_{hg} – number of pixels in the test sample collected in map category C_h that intersect reference category C_g;

n_{hh} – number of pixels in the test sample collected in map category C_h that intersect reference category $C_{g=h}$.

If a testing sample of n_h observations per mapped land cover class C_h, $h = 1,\ldots, M$, is randomly collected on the map and the reference land cover category C_g, $g = 1,\ldots, M$, is determined for each observation, then the estimated overall accuracy of the map is computed as

$$\hat{P} = \frac{1}{N}\sum_{h=1}^{M}\frac{N_h}{n_h}n_{hh} \tag{9.6}$$

and its respective estimated variance

$$\hat{V}(\hat{P}) = \sum_{h=1}^{M}\left(\frac{N_h}{N}\right)^2\frac{N_h - n_h}{N_h n_h}\left[\frac{n_{hh}}{n_h}\left(1-\frac{n_{hh}}{n_h}\right)\right]. \tag{9.7}$$

The estimated user's accuracy for each land cover class in the map is

$$\hat{P}_h = \frac{n_{hh}}{n_h} \tag{9.8}$$

and its respective estimated variance

$$\hat{V}(\hat{P}_h) = \frac{N_h - n_h}{N_h n_h}\left[\frac{n_{hh}}{n_h}\left(1 - \frac{n_{hh}}{n_h}\right)\right]. \tag{9.9}$$

The estimated producer's accuracy is the ratio of the estimates of N_{hh} and N_g,

$$\hat{P}_g = \frac{\hat{N}_{hh}}{\hat{N}_g} \tag{9.10}$$

and its respective estimated Mean Square Error (MSE)

$$\begin{aligned} \hat{MSE}(\hat{P}_g) = &\left(\frac{N}{\hat{N}_g}\right)^2 \sum_{h \neq g}\left[\left(\frac{N_h}{N}\right)^2\left(\frac{N_h - n_h}{N_h n_h}\right)\frac{n_{hg}}{n_h}\left(1 - \frac{n_{hg}}{n_h}\right)\hat{P}_g^2\right] + \\ &+\left[\left(\frac{N}{\hat{N}_g}\right)^2\left(\frac{N_h}{N}\right)^2\left(\frac{N_h - n_h}{N_h n_h}\right)\frac{n_{hh}}{n_h}\left(1 - \frac{n_{hh}}{n_h}\right)\left(1 - \hat{P}_g\right)^2\right] \end{aligned} \tag{9.11}$$

The unbiased estimates of N_{hh} and N_g can be calculated as

$$\hat{N}_{hh} = \frac{N_h}{n_h} n_{hh} \tag{9.12}$$

$$\hat{N}_g = \sum_{h=1}^{M} \frac{N_h}{n_h} n_{hg} \tag{9.13}$$

and their respective estimated variances

$$\hat{V}(\hat{N}_{hh}) = {N_h}^2 \frac{N_h - n_h}{N_h n_h}\left[\frac{n_{hh}}{n_h}\left(1 - \frac{n_{hh}}{n_h}\right)\right] \tag{9.14}$$

$$\hat{V}(\hat{N}_g) = \sum_{h=1}^{M} {N_h}^2 \frac{N_h - n_h}{N_h n_h}\left[\frac{n_{hg}}{n_h}\left(1 - \frac{n_{hg}}{n_h}\right)\right]. \tag{9.15}$$

A confidence interval for each P, at the approximate level 1-α, is computed as $\hat{P} \pm z_{1-\alpha/2}\left[\hat{V}(\hat{P})\right]^{1/2}$, where $z_{1-\alpha/2}$ is the constant exceeded with probability 1-α/2 by the N (0, 1) random variable. In accordance, the absolute precision $d(\hat{p})$ of each estimate $\hat{P}$ is computed as $Z_{1-\alpha/2}\left[\hat{V}(\hat{P})\right]^{1/2}$.

9.6. Results and Discussion

In Table 9.4 we introduce the percentage area occupied per mapped land cover class (N_h), and the 95% confidence intervals for the percentage area occupied per reference land cover class (N_g), overall accuracy (P), user's accuracy (P_h) and producer's accuracy (P_g). Overall accuracy of the final map was estimated at 80% with an absolute precision of 2% at the 95% confidence level. Estimated user's accuracies were, for almost all classes, superior to 60% and the respective absolute precisions estimates around 10% at the 95% confidence level. Point estimated producer's accuracies vary between 25 and 85%, while estimated absolute precisions range from 10 to 30% at the 95% confidence level. In Figure 9.1 we present the land cover map of Continental Portugal derived from multitemporal MERIS FR images for 2005.

Table 9.4. Percentage area occupied per mapped land cover class (N_h), and 95% confidence intervals for percentage area occupied per reference land cover class (N_g), overall accuracy (P), user's accuracy (P_h) and producer's accuracy (P_g). Absolute precision is represented by d.

Land cover class	N_h	$N_g \in \left[\hat{N}_g - d(\hat{N}_g), \hat{N}_g + d(\hat{N}_g)\right]$	$P_h \in \left[\hat{P}_h - d(\hat{P}_h), \hat{P}_h + d(\hat{P}_h)\right]$	$P_g \in \left[\hat{P}_g - d(\hat{P}_g), \hat{P}_g + d(\hat{P}_g)\right]$
5	0.8	[0.5, 0.9]	[0.7, 0.9]	[0.6, 1.0]
6	5.3	[3.3, 4.7]	[0.5, 0.7]	[0.6, 1.0]
7	0.5	[0.5, 0.7]	[1.0]	[0.7, 0.9]
11	0.3	[0.4, 0.8]	[0.8, 1.0]	[0.3, 0.7]
12	3.9	[3.5, 4.2]	[0.7, 0.9]	[0.7, 1.0]
21	12.9	[13.6, 13.8]	[0.9, 1.0]	[0.7, 0.9]
22	0.7	[1.8, 3.2]	[0.8, 1.0]	[0.1, 0.4]
23	0.3	[0.6, 1.6]	[0.8, 1.0]	[0.1, 0.4]
34	17.8	[16.0, 18.6]	[0.8]	[0.7, 1.0]
35	18.3	[16.0, 19.0]	[0.6, 0.8]	[0.5, 1.0]
38	2.8	[2.2, 3.2]	[0.6, 0.8]	[0.5, 1.0]
242	4.2	[2.5, 3.9]	[0.5, 0.7]	[0.5, 1.0]
311	7.1	[5.7, 7.7]	[0.6, 0.8]	[0.5, 1.0]
312	15.2	[12.2, 15.0]	[0.6, 0.8]	[0.6, 1.0]
321	3.5	[4.6, 6.2]	[0.8, 0.9]	[0.3, 0.8]
331	6.6	[4.9, 6.5]	[0.6, 0.8]	[0.6, 1.0]
Overall accuracy		$P \in \left[\hat{P} - d(\hat{P}), \hat{P} + d(\hat{P})\right]$	[0.78, 0.82]	

In Table 9.5 we introduce the proportion of matches between map and reference classes as a percentage of the total study area. Water (7) and Wetlands (5), Artificial areas (11 and 12), and Agricultural areas (21, 22, and 23) are the classes that show smaller out-diagonal row values. Indeed, given the specific spectral profiles of these classes along one year period, it was reasonably expectable that using temporal information in the classification feature space

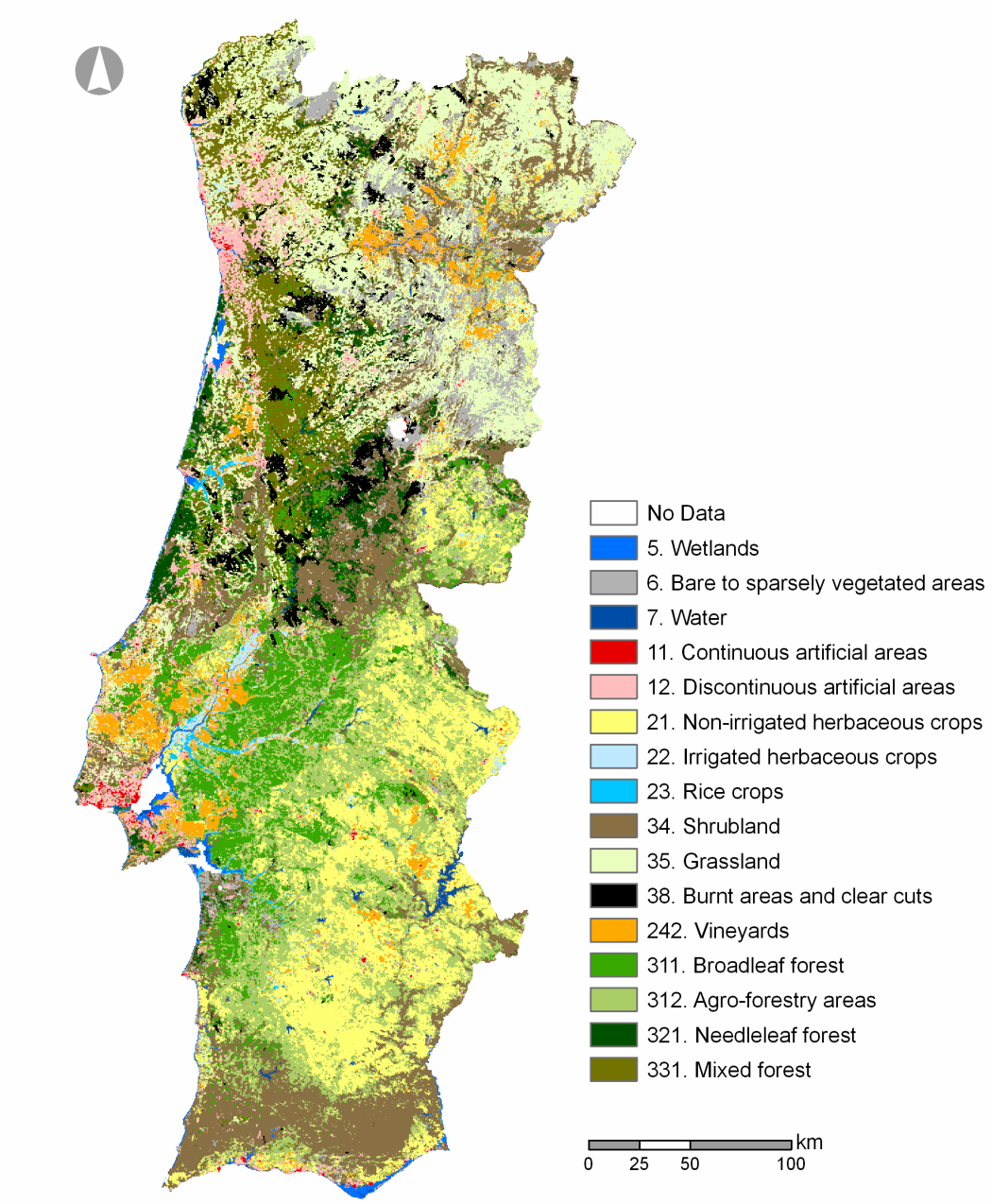

Figure 9.1. Land cover map for Continental Portugal derived from multitemporal MERIS images of 2005.

would improve their particular classification performances. Looking at per class confusions from Table 9.5, we perceive that natural land cover types show the higher out-diagonal row values that largely contribute to the overall class loss. Two different mismatch situations presented in Table 9.5 must be distinctively reported: (1) map classes that get confused with similar reference land cover classes; and (2) prohibited land cover confusions, i.e. between

dissimilar land cover classes. Regarding the first situation, we stress, for example, the mix between Bare to sparsely vegetated areas (6) and Grasslands (35), as well as the mix between Mixed forest (331) and Needleleaved forest (321). These confusions occur because these classes are constituted by similar land cover features that only vary upon different proportion arrangements within a pixel area. These classes represent the same land cover features, but in different proportions within a pixel area. In addition, these classes have contiguous geographical distributions and coexist in transitional land cover regions in the Portuguese landscape, which reinforce their mix. However, due to the similar classes' conceptual definition, their mix is not really problematic from a user's perspective.

Table 9.5. Estimated proportion ($\hat{N}_{hg}$) of matches between map and reference classes as a percentage of the total study area.

		Reference land cover class (C_g)															
		5	**6**	**7**	**11**	**12**	**21**	**22**	**23**	**34**	**35**	**38**	**242**	**311**	**312**	**321**	**331**
Map land cover class (C_h)	**5**	0.6		0.1													
	6		3.3			0.1	0.3	0.1		0.3	0.9	0.1	0.1	0.1	0.1	0.1	
	7			0.5													
	11				0.3												
	12				0.3	3.1		0.1			0.1			0.1	0.1		
	21						11.8	0.2	0.2	0.1			0.1		0.5		
	22							0.6	0.1								
	23								0.3								
	34	0.1	0.2				0.2	0.1		14.0	0.7	0.2		0.3	0.9	0.8	0.4
	35	0.1	0.2			0.5	0.7	0.8		1.1	12.5	0.3	0.3	0.4	0.6	0.4	0.5
	38		0.2							0.1	0.3	2.1		0.1		0.1	
	242		0.1				0.3	0.2	0.4	0.1	0.2		2.5	0.1	0.3		0.1
	311								0.1	0.6	0.3		0.1	5.0	0.4	0.5	0.2
	312						1.3	0.2	0.1	0.3	2.2		0.1	0.4	10.6		
	321									0.3	0.2	0.1				2.9	
	331		0.1				0.1	0.2		0.4	0.3			0.3	0.2	0.6	4.5

Looking at prohibited land cover confusions, we emphasize the mix between Shrubland (34) and Needleleaved forest (321), and also the mix between Broadleaf (311) and Needleleaved (321) forests. According to Caetano (2000), Shrublands (34) and forests get confused because the existence of shrubs under open forest canopies typically provokes forest to reflect similar to Shrublands (34). The confusion between Broadleaf and Needleleaved forests (311 and 321) is not so easily explained. Our guess is that some of the forested reference pixels may contain unidentified trees from several species. Therefore, they may be contaminating the spectral reflectances of testing pixels identified as single forest species. Indeed, it is sometimes difficult to identify the different tree species within forested pixels by visual analysis of aerial images, hence the existence of some errors in the reference database.

The analysis of the estimated producer's accuracies (Table 9.4) shows that there is a difficulty in mapping some land cover types, such as Artificial continuous areas (11) or Irrigated herbaceous crops (22). Still, these classes are mainly mapped as similar land cover classes. For example, the first is being extensively mapped as Artificial discontinuous areas (12). This confusion was somewhat expected because the land use is the same, but land cover differs within pixels according to the proportion of impermeable areas. Irrigated herbaceous crops (22) are being largely mapped as Grasslands (35). This confusion is due to the fact that some of Grasslands (35) individuals are irrigated during summer for livestock feeding and have intra-annual phenological trends identical to Irrigated herbaceous crops (22).

We present now a detailed comparison between our map and the CLC map for 2000. In Table 9.6 we depict per class intersections between the two products, as a percentage of the total study area. Intersections smaller than 0.1% were discarded from this comparison, thus contributing to an omission of approximately 2% of the total study area.

From the analysis of Table 9.6 we can state that our map is in accordance with the CLC2000 map. Minor disagreements are due to differences in the minimum mapping units (MMU) of products, while major disagreements are essentially due to land cover dynamics between 2000 and 2005. CLC2000 has a MMU of 25 ha and, as such, most polygons are not homogeneous because they result from manual aggregation and generalization of smaller polygons of different land covers. Moreover, CLC2000 is a land use-land cover map (LULC), whereas our product is merely a land cover map. Hence, a single land cover class in our map can have different uses properly depicted in CLC2000 map. For example, CLC2000 classes 332 (Bare rock), 333 (Sparsely vegetated areas), 331 (Beaches, dunes, sand plains) and 131 (Mineral extraction sites) correspond to a single LANDEO class (Bare to sparsely vegetated areas (6)). In this sense, we highlighted in Table 9.6 the cells that correspond to an obvious and admitted agreement between the two maps and nomenclatures. From this agreement, we show that at least 75% of the land cover in the territory is similarly represented by both cartographies. Note that this value does not include possible land cover dynamics between the two dates, such as forest becoming into shrubland, but only stationary land cover areas. Indeed, per class agreement analysis shows that major disparities between the two maps correspond to Shrublands (34), Grasslands (35) and Bare to sparsely vegetated areas (6).

Table 9.6. Comparison between land cover map derived from 2005 multitemporal MERIS images and CLC2000. Each cell represents the classes' intersection (%) in relation to the total area of the territory. Intersections with correct agreement between the two nomenclatures are in bold face.

		Classes from MERIS derived map for 2005															
		5	6	7	11	12	21	22	23	34	35	38	242	311	312	321	331
Classes from CLC map for 2000 for the Portuguese territory	**111**				**0.1**	**0.1**											
	112				**0.1**	**1.3**					**0.2**						
	121				**0.1**	**0.2**											
	122																
	123																
	124																
	131		**0.1**														
	132																
	133																
	141																
	142																
	211		0.2			**0.2**	**7.0**			0.3	2.0		0.3	0.1	**2.0**		0.1
	212					**0.1**	0.6	**0.4**			0.5		0.2		0.2		
	213							0.1	**0.3**				0.1		0.1		
	221		0.1			**0.1**	0.1			0.1	0.6		**1.5**		0.1		
	222						0.1			0.2	0.3		**0.2**		0.2		
	223		0.1			**0.1**	0.7			0.4	0.4		0.3	0.1	**1.1**		
	231						0.1				**0.2**						
	241		0.1			**0.4**	**0.6**			**0.4**	**2.2**		**0.1**		**0.4**		0.4
	242		0.2			**0.5**	**0.6**	**0.1**		**0.6**	**2.9**	0.1	**0.9**	0.1	**0.8**		0.2
	243		0.4			**0.3**	**0.6**			**1.7**	**2.6**	**0.1**	**0.2**	**0.2**	**1.0**	**0.1**	**0.4**
	244						**1.8**			**0.3**	**0.1**			**0.3**	**3.7**		
	311		0.2			**0.1**	0.3			2.9	0.6	**0.3**	0.1	**4.1**	**4.1**	0.1	**1.0**
	312		0.3			**0.1**				2.2	0.9	**0.5**		**0.3**		**2.1**	**1.2**
	313		0.1			**0.2**				0.8	0.8	**0.4**		**0.9**	**0.2**	**0.3**	**2.1**
	321		0.6				0.1			**0.5**	**0.6**	**0.1**			**0.1**		
	322		0.7				0.1			**1.3**	**1.0**	**0.3**			**0.2**	0.1	0.1
	323		0.1				0.1			**1.7**					**0.2**		
	324		1.0			**0.2**	0.2			**3.7**	**1.9**	**0.8**	0.1	**0.8**	**0.6**	**0.7**	**0.8**
	331		**0.1**														
	332		**0.3**								**0.1**						
	333		**0.4**							0.1	**0.2**	**0.1**					
	334		0.1							0.1	**0.1**						
	411																
	421	**0.2**															
	422	**0.1**															
	423																
	511	**0.1**								0.1							
	512	**0.1**		**0.1**						0.1						0.1	
	521																
	522	**0.1**		**0.1**													
	523	**0.1**		**0.1**													
Overall agreement		0.7	0.9	0.3	0.3	3.9	10.6	0.5	0.3	10.2	12.1	2.6	2.9	6.6	14.4	3.2	5.5
N_h **(Table 9.4)**		0.8	5.3	0.5	0.3	3.9	12.9	0.7	0.3	17.8	18.3	2.8	4.2	7.1	15.2	3.5	6.6

Let us now compare our classification results with those from other maps recently derived for particular areas of interest from automatic classification of MERIS FR images. Clevers *et al.* (2007) produced a map for the Netherlands with seven relevant land cover classes for

2003 with multitemporal MERIS FR data. Classification results were evaluated by comparing their map with the Dutch land use database (LGN) – a map with 25 m cell size derived from automatic classification of higher spatial resolution satellite images. The match between both maps was of 67%. García-Gigorro *et al.* (2007) have presented specific land use-land cover (LULC) maps for two Spanish provinces using MERIS FR images from a single date. The nomenclature of these maps was derived using the LCCS of FAO and consists of eleven land cover types similar to the nomenclature of our final map. Overall accuracies of presented maps were of approximately 74%. The estimated accuracy of our map follows that attained in previous studies, hence reinforcing the aptitude of land cover products derived from medium spatial resolution satellite images for national applications. However, since our map was assessed through an independent test sample, collected through a stratified random sampling design, we statistically demonstrated the assertions from previous studies, which were simply based upon confidence-building approaches and/or unseen training sites.

9.7. Conclusion

This study confirmed that land cover maps derived from automatic classification of medium spatial resolution satellite images are suitable for national applications. To assert this, we produced a land cover map of Continental Portugal from MERIS FR multitemporal images of 2005, evaluated its overall and per class accuracy, and compared it with a reliable land cover map derived for the Portuguese territory from visual analysis of high resolution satellite data. In the sequence, we followed a methodological approach that involved:

(1) The definition of an exhaustive land cover nomenclature with 16 land cover classes that comprise the most important landscape characteristics of the study area – derived through the LCCS of FAO;

(2) The use of distinct classification strategies for intra-annual stationary and transient land cover classes' characterization – these strategies included, respectively, a Linear Discriminant Classifier (LDC) and a vegetation differentiating technique;

(3) The accuracy assessment of the final map using an independent testing sample collected through a stratified random sampling design.

An 80% overall accuracy for the final land cover map was estimated with an absolute precision of 2% at the 95% confidence level. This result statistically confirmed the assertions from previous studies on automatic land cover classification from MERIS FR data for particular areas of interest, e.g. Clevers *et al.* (2007) and García-Gigorro *et al.* (2007), to cite

but a few. Indeed, our estimated overall accuracy corroborated that medium spatial resolution satellite images can be used to automatically derive consistent land cover maps comparable to those with exhaustive nomenclatures, which are regularly used for national applications. On the other hand, estimated overall accuracy of our land cover map largely exceeds the estimated overall accuracy of global land cover products derived from automatic classification of satellite images with similar spectral, temporal and spatial resolutions for the same study area. We showed that accuracy discrepancies between our map and those global maps for the Portuguese territory are due to methodological constrains, namely the ontological definition of land cover classes and adopted classification strategies, and not due to satellite images characteristics. Overall and per class accuracy estimates for the final map confirmed that our methodological approach proved successful in discriminating between the land cover classes that better characterize the landscape of the study area at sensor's spatial resolution, whereas global land cover products do not provide so reliable information.

9.8. Acknowledgements

This study was carried out in the framework of the project "LANDEO - User driven land cover characterisation for multi-scale environmental monitoring using multi-sensor earth observation data (PDCTE/MGS/49969/2003)" funded by "Programa Dinamizador das Ciências e Tecnologias para o Espaço" from "Fundação para a Ciência e Tecnologia", and from the "Announcement of Opportunity for the Utilisation of ERS and ENVISAT Data" from European Space Agency (ESA). Research by Hugo Carrão was founded by the "Fundação para a Ciência e Tecnologia" (SFRH/BD/18447/2004).

9.9. References

Arino, O., Leroy, M., Ranera, F., Gross, D., Bicheron, P., Nino, F., Brockman, C., Defourny, P., Vancutsem, C., Achard, F., Durieux, L., Bourg, L., Latham, J., Di Gregorio, A., Witt, R., Herold, M., Sambale, J., Plummer, S., Weber, J-L., Goryl, P. and Houghton, N., 2007, GLOBCOVER - A Global Land Cover Service with MERIS. In *Proceedings of the ENVISAT Symposium* (unpaginated CD-ROM), 23-27 April 2007, Montreaux, Switzerland.

Armas, R. and Caetano, M., 2005, Scale effect on burned area assessment using earth observation data. In M. Oluic (Ed.), *New Strategies for European Remote Sensing* (pp. 61-67). Rotterdam: Millpress.

Ballantine, J.A.C., Okin, G.S., Prentiss, D.E. and Roberts, D.A., 2005, Mapping North African landforms using continental scale unmixing of MODIS imagery. *Remote Sensing of Environment*, 97, 470-483.

Bartholomé, E. and Belward, A.S., 2005, GLC2000: a new approach to global land cover from Earth observation data. *International Journal of Remote Sensing*, 26, pp. 1959-1977.

Belward, A.S., Estes, J.E. and Kline, K.D., 1999, The IGBP-DIS global 1-km land-cover data set DISCover: a project overview. *Photogrammetric Engineering and Remote Sensing*, 65, 1013–1020.

Borak, J.S. and Strahler, A.H., 1999, Feature selection and land cover classification of a MODIS-like data set for a semiarid environment. *International Journal of Remote Sensing*, 20, 919-938.

Caetano, M., 2000, Characterization of coniferous forest understory with remotely sensed data. Lisboa, Portugal: Universidade Técnica de Lisboa, 318 pp.

Caetano, M., Carrão, H. and Painho, M., 2005, *Alterações da ocupação do solo em Portugal Continental: 1985 – 2000*. Amadora, Portugal: Instituto do Ambiente, 52 pp.

Caetano, M. and Araújo, A., 2006, Comparing land cover products CLC2000 and MOD12Q1 for Portugal. In A. Marçal (Ed.), *Global Developments in Environmental Earth Observation from Space* (pp. 469-477). Rotterdam: Millpress.

Card, D. (1982), Using known map category marginal frequencies to improve estimates of thematic map accuracy. *Photogrammetric Engineering and Remote Sensing*, 48, 431–439.

Carrão, H., Gonçalves, P. and Caetano, M., 2006a, MERIS Based Land Cover Characterization: A Comparative Study. In *Proceedings of the ASPRS 2006 Annual Conference - Prospecting for Geospatial Information Integration* (unpaginated CD-ROM), 1-6 May 2006, Reno, Nevada, USA.

Carrão, H., Capão, L., Bação, F. and Caetano, M., 2006b, MERIS Based Land Cover Classification With Self-Organizing Maps: Preliminary Results. In Proc*eedings of the 2nd EARSeL SIG Workshop on Land Use & Land Cover* (unpaginated CD-ROM), 28 - 30 September 2006, Bonn, Germany.

Carrão, H., Araújo, A., Cerdeira, C., Sarmento, P., Capão, L. and Caetano, M., 2007a, A reference sample database for the accuracy assessment of medium spatial resolution

land cover products in Portugal. In *Proceedings of the IEEE International Conference on Geoscience and Remote Sensing Symposium* (IGARSS´2007), 20-29 July 2007, Barcelona, Spain.

Carrão, H., Gonçalves, P. and Caetano, M., 2007b, Contribution of multispectral and multitemporal information from MODIS images to land cover classification. *Remote Sensing of Environment*, 112, 986-997.

Carrão, H., Gonçalves, P. and Caetano, M., 2007c, Use of intra-annual satellite imagery time-series for land cover characterization purpose. *EARSeL eProceedings*, 6, 1-11.

Carrão, H., Caetano, M. and Coelho, P., 2007d, Sample Design and Analysis for Thematic Map Accuracy Assessment: An Approach Based on Domain Estimation for the Validation of Land Cover Products. In *Proceedings of the 32nd International Symposium on Remote Sensing of Environment* (unpaginated CD-ROM), 25-29 June 2007, San Jose, Costa Rica.

Cerdeira, C., Araújo, A., Carrão, H. and Caetano, M., 2006, Validação das Cartografias Globais de Ocupação do Solo, GLC2000 e MOD12Q1, para Portugal Continental. In *Actas do IX Encontro de Utilizadores de Informação Geográfica* (unpaginated CD-ROM), 15 – 17 November 2006, Oeiras, Portugal.

Cihlar, J., Ly, H. and Xiao, Q., 1996, Land cover classification with AVHRR multichannel composites in northern environments. *Remote Sensing of Environment*, 58, 36-51.

Clevers, J. G. P. W., Schaepman, M. E., Mücher, C. A., de Wit, A. J. W., Zurita-Milla, R. and Bartholomeus, H. M., 2007, Using MERIS on Envisat for land cover mapping in the Netherlands. *International Journal of Remote Sensing*, 28, 637-652.

Cochran, W., 1977, *Sampling Techniques*, 3rd Ed. New York: Wiley, 428 pp.

Curran, P. and Steele, C., 2005, MERIS: the re-branding of an ocean sensor. *International Journal of Remote Sensing*, 26, 1781-1798.

Dash, J., Mathur, A., Foody, G.M., Curran, P.J., Chipman, J.W. and Lillesand, T.M., 2007, Land cover classification using multi-temporal MERIS vegetation indices. *International Journal of Remote Sensing*, 28, 1137-1159.

DeFries, R., Hansen, M., Townshend, J.R.G. and Sholberg, R., 1998, Global land cover classifications at 8km spatial resolution: The use of training data derived from Landsat

imagery in decision tree classifiers. *International Journal of Remote Sensing*, 19, 3141-3168.

DeFries, R. and Belward, A.S., 2000, Global and regional land cover characterization from satellite data: an introduction to the Special Issue. *International Journal of Remote Sensing*, 21, 1083-1092.

DGRF - Direcção Geral dos Recursos Florestais, 2005, Inventário Florestal Nacional. Available online at: http://www.dgrf.min-agricultura.pt/ifn/index.html (accessed 28 June 2005).

Di Gregorio, A. and Jansen, L.J.M., 2000, *Land Cover Classification System*. Rome, Italy: FAO, 179 pp.

EEA - European Environment Agency, 2002, I&CLC2000 Technical Reference Document. Joint Research Centre, European Comission, Available online at: http://terrestrial.eionet.eu.int.

Friedl, M.A., McIver D.K., Hodges J.C.F., Zhang X.Y., Muchoney D., Strahler A.H., Woodcock, C.E., Schneider A., Cooper, A., Baccini, A., Gao, F. and Schaaf, C., 2002, Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment*, 83, 287-302.

García-Gigorro, S., González-Alonso, F., Merino-de-Miguel, S., Roldán-Zamarrón, A. and Cuevas, J. M., 2007, MERIS-FR potential for land use-land cover mapping in Spain. *International Journal of Remote Sensing*, 28, 1405-1412.

Giri, Z., Zhu, Z. and Reed, B., 2005, A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets. *Remote Sensing of Environment*, 94, 123-132.

Hammond, T. O. and Verbyla, D. L., 1996, Optimistic bias in classification accuracy assessment. *International Journal of Remote Sensing*, 17, 1261-1266.

Hansen, M.C., DeFries, R.S., Townshend, J.R.G. and Sohlberg, R., 2000, Global land cover classification at 1km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, 21, 1331-1364.

Hodges, J., 2001, MODIS MOD12 Land Cover and Land Cover Dynamics Products User Guide. Available at: http://geography.bu.edu/landcover/userguidelc/intro.html (last accessed, 3 June, 2005).

Jensen, J.R., 1996, *Introductory Digital Image Processing: A Remote Sensing Perspective*, 2nd Ed. New Jersey: Prentice Hall Inc., 379 pp.

Johnson, R.A. and Wichern, D.W., 1998, *Applied Multivariate Statistical Analysis*, 4th Ed. New Jersey: Prentice Hall, Upper Saddle River, 767 pp.

Kuncheva, L., 2004, *Combining Pattern Classifiers: Methods and Algorithms*. New Jersey: Wiley-Interscience, 350 pp.

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

Matsuoka, M., Hayasaka, T., Fukushima, Y. and Honda, Y., 2004, Land Cover Classification Over Yellow River Basin Using Satellite Data. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium* (IGARSS), 20-24 September, Fairbanks, USA.

Salvador, R., Valeriano, J., Pons, X. and Díaz-Delgado, R., 2000, A semi-automatic methodology to detect fire scars in shrubs and evergreen forests with Landsat MSS time series. *International Journal of Remote Sensing*, 21, 655-671.

Stehman, S. V., 2001, Statistical rigour and practical utility in thematic map accuracy. *Photogrammetric Engineering and Remote Sensing*, 67, 727–734.

Strahler, A., 2002, Global Land Cover: Approaches to Validation. GLC2000 Meeting, JRC, Ispra, March 2002. Available online at: http://www-gvm.jrc.it/glc2000/Workshops/March2002/Presentations/strahler_Validation.pps (accessed 15 January 2007).

Wessels, K.J., De Fries, R.S., Dempewolf, J., Anderson, L.O., Hansen, A.J., Powel, S.L. and Moran, E.F., 2004, Mapping regional land cover with MODIS data for biological conservation:Examples from the Great Yellowstone Ecosystem, USA and Pará State, Brazil. *Remote Sensing of Environment*, 92, 67-83.

Wickham, J.D., Stehman, S.V., Smith, J.H. and Yang, L., 2004, Thematic accuracy of the 1992 National Land-Cover Data for the western United States. *Remote Sensing of Environment*, 91, 452–468.

Wilkinson, G., 2005, Results and Implications of a Study of Fifteen Years of Satellite image classification Experiments. *IEEE Transactions on Geoscience and Remote Sensing*, 43, 433-440.

Wulder, M., Franklin, S., White, J., Link, J. and Magnussen, S., 2006, An accuracy assessment framework for large-area land cover classification products derived from

medium-resolution satellite data. *International Journal of Remote Sensing*, 27, 663–683.

Zhu, Z., Yang, L., Stehman, S.V. and Czaplewski, R.L., 2000, Accuracy assessment from the US Geological Survey Regional Land Cover Mapping Program: New York and New Jersey Region. *Photogrammetric Engineering and Remote Sensing*, 66, 1425–1435.

10 A NONLINEAR HARMONIC MODEL FOR FITTING SATELLITE IMAGES TIME SERIES: ANALYSIS AND PREDICTION OF LAND COVER DYNAMICS

Carrão, H., Gonçalves, P. and Caetano, M. (2009). A nonlinear harmonic model for fitting satellite images time series: analysis and prediction of land cover dynamics. *IEEE Transactions on Geoscience and Remote Sensing*, submitted.

Abstract

Numerous efforts have been made to develop models to fit multispectral reflectances and vegetation indices (VI) time series from satellite images for diverse land cover classes. The common objective of these models is to derive a set of measurable parameters able to characterize and to reproduce the land cover dynamics of natural- and human-induced ecosystems. Good-fitting models should therefore match different waveforms and be insensitive to sharp and localized variations, generally due to atmospheric disturbances. In this paper we propose a model-based approach to identify and predict important dynamics for indiscriminate land cover classes. Our method relies on an original nonlinear harmonic model that remarkably matches intra-annual time series of multispectral reflectances and VIs obtained from satellite images. The proposed model is (i) parsimonious, comprising only 5 parameters; (ii) readily identifiable (in the maximum likelihood sense) from only few observations; (iii) robust to noise; and (iv) versatile, since it can reproduce a wide variety of intra-annual land cover dynamics as a deterministic function of time. To demonstrate the relevance of our approach, we use a time series of MODIS 8-day composite images acquired in Portugal over a one-year period at a 500 m nominal spatial resolution. For 13 different land cover classes, representative of Mediterranean landscapes, we evaluate the data-model adequacy of our model and compare it with several other approaches. We then address a particularly interesting and promising application of our method using rice crops and shrublands as case studies. We not only show that phenological attributes can be accurately estimated from the fitted time series, but we also demonstrate that it is possible to make early predictions of phenological attributes dates and magnitudes from our expected model adjusted to only few anterior observations.

Index Terms: Curve fitting, intra-annual dynamics, land cover, nonlinear harmonic model, phenology, prediction, remote sensing, time series.

10.1. Introduction

The spectral, temporal and spatial information contained in satellite images time series offers unique opportunities for monitoring and managing land cover dynamics from local to global spatial scales [1]. Because satellite images time series provide temporal and spatial consistent spectral measurements of Earth's surface that correlate directly with annual cycles of vegetation growth, i.e. phenology, they offer a significant insight into the response of vegetation to short-term environmental forcing effects, both natural and anthropogenic [2], [3]. Nonetheless, to effectively track continuous intra-annual dynamics of land surface, such as vegetation phenology, error-free remotely-sensed data acquired at appropriate time instants are stringent conditions rarely met in practice. Indeed, satellite images time series correspond to discrete, possibly non-uniformly sampled measurements of spectral variations, which moreover, can be corrupted by atmospheric, geometric and radiometric disturbances [4]. Therefore, to reduce imprecision on phenological attributes estimation, we must resort to denoising procedures and fill the gaps between satellite images acquisition dates with a pertinent continuous interpolating function [4]. Regarding noise, smoothing techniques (e.g. running averages, running medians or compound smoothers) are commonly used to suppress local artifacts in time series [5], [6]. As for interpolation, more sophisticated mathematical models make it possible to estimate phenological attributes from the functional analysis of continuous time evolutions of land cover spectral responses, e.g. [3], [5], [7]-[10]. Indeed, these functions increase the spatial and temporal consistency of satellite images time series used to derive phenological attributes [11]. Nevertheless, remote sensing estimations of phenological markers are available only at images acquisition dates [4], thus delaying the identification of land cover dynamics and the mitigation of negative stresses on the environment.

In this paper we propose a model-based approach for the identification and prediction of phenological attributes from satellite images time series. Our main goal is to derive phenological attributes in early time to serve as input for decision systems (e.g. forest fire prevention and mitigation operational systems, such as WFAS [12] in the US and PREMFIRE [13] in the EU) and support policy makers within their decision-making frameworks (e.g. agricultural programs, such as the Monitoring Agriculture with Remote Sensing (MARS) programme of the EU [14] and the Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing (AgRISTARS) in the US [15]). We believe this is an innovative approach: while previous studies have focused only on the identification of phenological key-attributes by fitting the entire annual time series, what we address here is the early prediction of phenological attributes from few anterior observations. Our method

relies on an original parsimonious nonlinear harmonic model that we previously proposed [16] to match intra-annual time series of multispectral reflectances and VIs obtained from satellite images. The paper is organized as follows: in Section 10.2, after listing most standard models for fitting satellite images time series, we present the analytic form of our nonlinear model and discuss its identifiability properties. In Section 10.3, we test the robustness and the flexibility of this parametric model on MODIS time series corresponding to multispectral reflectance and VIs measurements acquired in Portugal for 13 land cover classes with distinct intra-annual variations. In this Section, we also focus on a particularly seducing application. From the modelled time series, we identify and predict phenological key-markers that are highly relevant for the underlying land cover classes. We end in Section 10.4, with concluding remarks and future work.

10.2. Satellite images time series fitting

Before elaborating on our original model, we start this Section with a brief survey of some of the most standard approaches currently used for modelling satellite images time series. We stress the main advantages and weaknesses of each of these models.

10.2.1. State-of-the-art

Because annual phenologies of vegetation have seasonal paths, simple harmonic oscillators and Fourier-series are among the principal models used to fit satellite image time series [3], [10]. However, while simple harmonic oscillators fail at reproducing nonlinear waveforms, Fourier series approaches implicitly assume regularly spaced time series and underlying components that are truly harmonic [3], [5], [11]. Thus, these functions are too constrained to satisfactorily model multispectral reflectances for a wide range of land cover classes. To overcome these difficulties, refined piecewise-defined functions were recently proposed for fitting land cover time series derived from satellite images.

References [5] and [17] presented a new method based on piecewise-defined local Asymmetric Gaussian (AG) functions for extracting seasonality information from NDVI time series data. In this algorithm, local model functions with seven parameters are fitted to data in intervals around maxima and minima in the time series and used to build a global function that describe the annual cycles of vegetation. In the same vein, [9], [11], and [18] proposed piecewise-defined local Double Logistic (DL) functions to identify land cover time trajectories from VIs time series. The local model functions rely only on six parameters, but have the same general form of the former and can also be merged to build a global function that describe the annual seasons of land cover types. Generically, the global AG and DL models functions have the form:

$$f(t) = \sum_{h=1}^{H} c_{1,h} + c_{2,h} g(t; A_h) \tag{10.1}$$

where H designates the number of local model functions, i.e. the number of extrema values (minima and maxima) of the time series to be modelled. The linear parameters $c_{1,h}$ and $c_{2,h}$ determine, respectively, the base level and the amplitude of the local function h, whereas the nonlinear meta-parameter $A_h = (a_1,\ldots,a_r)$, determines its shape. The AG type basis function has the form [5]:

$$g(t; a_1,\ldots,a_5) = \begin{cases} \exp\left[-\left(\dfrac{t-a_1}{a_2}\right)^{a_3}\right] & \text{if} \quad t > a_1 \\ \exp\left[-\left(\dfrac{a_1-t}{a_4}\right)^{a_5}\right] & \text{if} \quad t < a_1 \end{cases} \tag{10.2}$$

In this function, a_1 determines the position of the maximum (or minimum) of g; a_2 and a_3 (respectively a_4 and a_5) determine the width and flatness of the right (respectively left) half of the function. Accordingly, the AG global model has a total of $H\times7$ parameters. Regarding the DL type basis function [11], it reads:

$$g(t; a_1,\ldots,a_4) = \frac{1}{1+\exp\left(\dfrac{a_1-t}{a_2}\right)} - \frac{1}{1+\exp\left(\dfrac{a_3-t}{a_4}\right)} \tag{10.3}$$

where a_1 and a_3 determine, respectively, the position of the left and right inflection points, and a_2 and a_4 fix the rates of change at those points. Therefore, the DL global function has a total of $H\times6$ parameters.

Albeit superior to standard harmonic analyses, these methods still suffer from severe constrains that hamper their effective use. In particular, these models are not totally versatile as they are always adopted to the upper envelope of the time series [5], [11]. Thus, the fits are subjective when there is no clear seasonality in land cover time series (e.g. urban areas, water or barren) or the time series being modeled are other than VIs. Moreover, growing season variations that do not coincide with the shape of the Gaussian or Logistic curves are not wholly captured [11]. AG and DL functions are not robust to noise and can only fit smoothed time series [5]. When the noise level is high (such as for the corresponding satellite image time series), these methods are generally not successful because there is a difficulty for estimating the time t_h corresponding to the occurrence dates of extrema values (minima and maxima) of the time series to be modelled. More importantly though, since three consecutive annual time series are required to properly estimate the intra-annual

functions corresponding to the central year [5], the model parameters are not easily identifiable. Finally, as local contributions of the time limited functions are non-adaptive, phenological attributes for specific land cover classes can not be early predicted.

10.2.2. Nonlinear harmonic model

Definition

In this study we recall the model we proposed in a previous work [16] to fit intra-annual response of land cover multispectral reflectances and VIs obtained from satellite image time series. It consists of a trigonometric parsimonious function that describes the nonlinear harmonic solution of a chaotic attractor [19]:

$$\{S_\theta(t) = M + A\cos(\omega_0 t + \phi + \alpha\cos(\omega_0 t + \varphi)),\quad \theta \in \Theta\} \tag{10.4}$$

In words, it is a sinusoidal parametric model with an *a priori* fixed annual frequency $\omega_0 = 2\pi.T^{-1}$, where T denotes a one-year period, and a nonlinear phase term $\alpha\cos(\omega_0 t + \varphi)$. The latter plays the role of a time warping (acceleration/deceleration) to account for the different regimes in the intra-annual cycles of land cover types (e.g. growths, cuts, harvests, droughts).

The meta-parameter $\theta = (M, A, \phi, \alpha, \varphi)$ belongs to the parameter space $\Theta = \{(M, A, \phi, \alpha, \varphi): M \in \Re, A \geq 0, \phi \in [0, 2\pi], \alpha \geq 0, \varphi \in [0, 2\pi]\}$, and each component has a specific action in the model response:

- M is a linear parameter that determines the base level of the modelled variable and represents its annual mean;
- A is the amplitude for the sine wave that fixes the peak deviation from the annual mean of the modelled variable;
- ϕ is the annual phase that allows to reproduce the specific seasoning of a given land cover class;
- α controls the nonlinear strength. When $\alpha = 0$, the model reduces to a linear harmonic oscillator, whereas $\alpha > 0$ introduces non-symmetry in the waveform;
- φ is the annual nonlinear phase that defines the type of non-symmetry of the waveform. This phase allows time to "slow down" and to "accelerate", in order to reproduce asymmetries in variations (increases *vs*. decreases) or in steady state durations.

Parameters identification

Given a discrete time series $\mathbf{x} = \{x(t_i): i = 1,\ldots, N;\ 1 \le t_i \le T\}$, where N is the number of observations, we want to find the "best" approximation $\mathbf{s}_{\boldsymbol{\theta}} = \{s_{\theta}(t_i): i = 1,\ldots, N;\ 1 \le t_i \le T\}$ in a maximum likelihood sense. To this end we assume the following observation model:

$$\mathbf{x} = \{x(t_i) = s_{\theta}(t_i) + \varepsilon(t_i): i = 1,\ldots, N;\ 1 \le t_i \le T\}, \qquad (10.5)$$

where $\boldsymbol{\varepsilon}$ is simplistic supposed to be a stationary zero-mean Gaussian noise of unknown variance σ_{ε}^2[3]. Under the hypothesis of independent noise observations, the likelihood function reads

$$f(\mathbf{x}|\ \mathbf{s}_{\theta}) = \left(2\pi\sigma_{\varepsilon}^2\right)^{-\frac{N}{2}} \exp\left\{-\frac{1}{2\sigma_{\varepsilon}^2}\sum_{i=1}^{N}\left(x(t_i)-s_{\theta}(t_i)\right)^2\right\}, \qquad (10.6)$$

and the maximum likelihood estimate $\hat{\theta}$ becomes

$$\hat{\theta} = \arg\min_{\theta\in\Theta}\sum_{i=1}^{N}\left(x(t_i) - s_{\theta}(t_i)\right)^2 . \qquad (10.7)$$

Formally, (10.7) is the solution of the system obtained by differentiating the l_2-norm $\|\mathbf{x} - \mathbf{s}_{\theta}\|^2$ with respect to each component of θ and equaling the results to zero. However, because of the trigonometric form defining (10.4), we are led to a system of nonlinear equations that we can not solve analytically. Considering well-posed problems that involve at least as many observations as unknowns, we then proposed to numerically minimize the approximation error $\|\mathbf{x} - \mathbf{s}_{\theta}\|^2$ using a nonlinear least-squares optimization routine based on the Levenberg-Marquardt method [20], [21]. Furthermore, as the problem may not be convex, in practice we initialize the solving routine with the heurist initial guess $\hat{\theta}_0$:

- $\hat{M}_0$ is set to the empirical median of $\mathbf{x}$, i.e. $(\max\{x(t_i)\} + \min\{x(t_i)\}) / 2$;
- $\hat{A}_0$ is set to the maximum deviation of $\mathbf{x}$, i.e. $(\max\{x(t_i)\} - \min\{x(t_i)\}) / 2$;
- $\hat{\phi}_0$, $\hat{\alpha}_0$, and $\hat{\varphi}_0$ are determined from a least-square regression of the nonlinear phase term $\delta(t_i) = \omega_0 t_i + \phi + \alpha\cos(\omega_0 t_i + \varphi)$ in (10.4).

Let us recall that we purposely consider a one-year periodic model. Shorter phenological periods, like the pseudo half-period of bi-annual crops, are accounted for by the non-linearity

[3] *Although simplistic and sub-optimal in practice, this Gaussian assumption allows for calculating the likelihood function and to express the MLE estimator of* $\hat{\theta}$ *in a close form.*

in $\delta(t_i)$. Therefore, to cope with intermediate phase shifts, we must properly unwrap the instantaneous phase $\hat{\delta}(t_i) = \cos^{-1}\left(\left(x(t_i) - \hat{M}_0\right) / \hat{A}_0\right)$. We then use a geometric criterion to keep phase continuity at global extrema, and to let the angle increase beyond $2k\pi$, (k = 1, 2,...), at local extrema. The parameter set $o = \left(\hat{\phi}_0, \hat{\alpha}_0, \hat{\varphi}_0\right)$ that belongs to the parameter space $\mathrm{O} = \left\{\left(\hat{\phi}_0, \hat{\alpha}_0, \hat{\varphi}_0\right) : \hat{\phi}_0 \in [0, 2\pi],\ \hat{\alpha}_0 \geq 0,\ \hat{\varphi}_0 \in [0, 2\pi]\right\}$ is finally initialized to the solution of the minimization problem: $\hat{o} = \arg\min_{o \in \mathrm{O}} \sum_{i=1}^{N} \left(\left(\hat{\delta}(t_i) - \omega_0 t_i\right) - \left(\hat{\phi}_0 + \hat{\alpha}_0 \cos\left(\omega_0 t_i + \hat{\varphi}_0\right)\right)\right)^2$.

Model identifiability

The model's identifiability and robustness address the following two questions: 1) is it possible to exactly retrieve the parameter values of a given model simply from a finite sampled sequence of its time response? 2) given a model response generated under controlled experimental conditions (notably regarding the sampling process and the Signal to Noise Ratio (SNR)), can the parameters of the model be estimated with acceptable precision? A standard practice for studying identifiability of a parametric model is then to consider multiple known input datasets which are then compared to the corresponding fitted models [22]. The input datasets are model outputs computed with *a priori* known parameters values. Thus, for different values of the meta-parameter θ of (10.4), test signals $s_\theta(t)$ were synthesized over a one year period (T = 365). Experimental conditions are summarized in Table 10.1.

Table 10.1. Model parameters used to synthesize time series data.

M	***A***	***α***	***ϕ (rad)***	***φ (rad)***
0	1	0:0.1:3	0: π/3:2π	0: π/3:2π

First, to assess model identifiability as function of N, we sampled the synthesized test signals $s_\theta(t)$ on regularly spaced time lattices with N = {5, 6, 8, 12, 16, 24, 46} observations. For each combination of N and the meta-parameter θ in Table I, we determine the maximum likelihood estimate $\hat{\theta}$ and compute the fitting error $e_N = \sqrt{\frac{\int_1^{365} \left(s_\theta(t) - \hat{s}_{\hat{\theta}}(t)\right)^2 dt}{365 - 1}} \times \frac{100}{P[s\theta]}$, expressed as a percentage, as a measure of agreement between our functional model and the synthesized signals. In Figure 10.1 we present the average fitting error, $< e_N >_\theta$, corresponding to the same N.

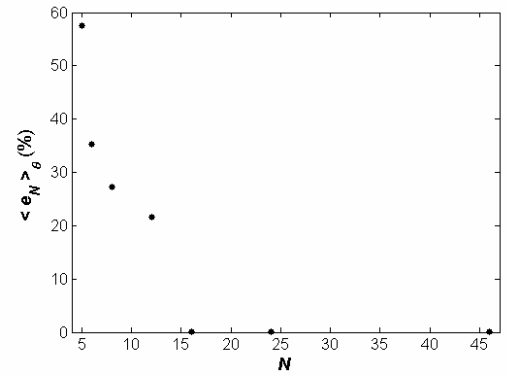

Figure 10.1. Average fitting error (%), $< e_N >_\theta$, corresponding to the same *N*.

From Figure 10.1, we see that the average fitting error decreases as long as the number of modelling observations increases and that, independently of their magnitude, the model parameters are perfectly estimated for $N \geq 16$. A closer view at the corresponding time series indicates that this result is not due to methodological constrains, but because of data restrictions. The model failed at matching the time series only when the quantization step of the synthesized signal was bigger than the respective intra-annual variations, i.e. the model parameters could not be perfectly estimated when the discrete time lattice was insufficient to properly describe the underlying continuous signal.

Next, to simulate different input SNR values, we sampled the synthesized signals on 8-day spaced time lattices, i.e. $N = 46$, and a stationary zero-mean white Gaussian noise $\varepsilon(t)$ with variable power σ_ε^2 is added to the time series,

$$SNR_{input} = 10 \log_{10} \frac{P[s_\theta]}{\sigma_\varepsilon^2}, \tag{10.8}$$

For each combination of the meta-parameter θ and of different input SNR values, we determine the maximum likelihood estimate $\hat{\theta}$ and define the output SNR as:

$$SNR_{output} = 10 \log_{10} \left(\frac{P[s_\theta]}{P[s_{\hat{\theta}} - s_\theta]} \right). \tag{10.9}$$

We choose to assess identifiability of the model through the SNR gain $\gamma = SNR_{output} - SNR_{input}$. Alternatively, we could also compare the exact value of the meta-parameter θ to its estimate $\hat{\theta}$. However this approach raises the issue of defining a consistent metric to evaluate

the distance between vectors whose components lie in different spaces. Figure 10.2 shows for four different input SNRs, examples of generated noisy time series superimposed to the corresponding fitted models.

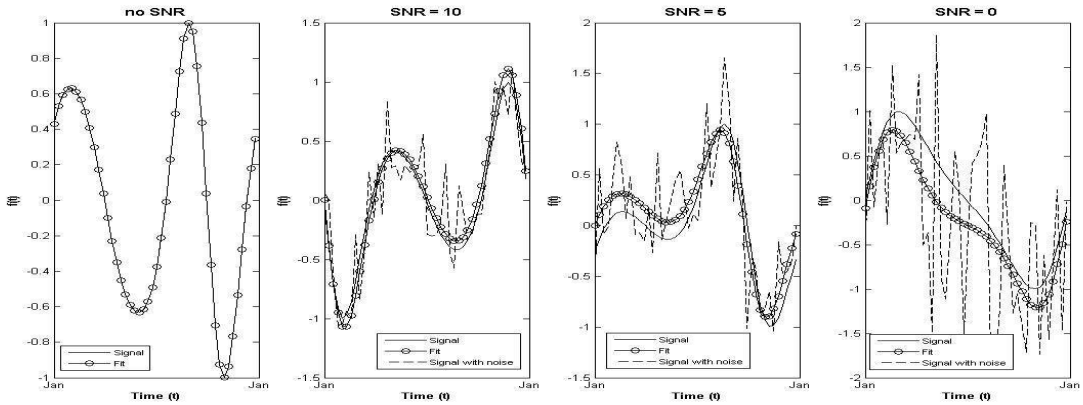

Figure 10.2. Nonlinear model fits for time series synthesized with different SNRs.

From our experiments, we first observed that for a given input SNR, the output SNR does not significantly depend on the values of θ. This led us to the interesting conclusion that the sensitivity of our maximum likelihood estimation procedure is independent of the magnitude of the model parameters. That is the reason why Table 10.2 presents only the averaged SNR gain, $<\gamma>_{\theta}$, corresponding to the same input SNRs. In all the cases, not only the underlying model parameters are reasonably estimated, but also the SNRs systematically increase by more than 10dB! Given also the relatively low computational cost of the method, a reduction by a factor 10 of the disturbance intensity, is a substantial improvement within the context of satellite imaging.

Table 10.2. Mean SNR and SNR gain for model fits to input time series with different SNRs.

Input time series SNR (dB)	**Mean SNR (dB) for model fits**	**Mean SNR (dB) gain for model fits**
-10	1.7	11.7
-5	6.8	11.8
0	11.6	11.6
5	17.1	12.1
10	22.9	12.9
∞	200.9	-

10.3. Case studies

In this section we introduce the proposed model-based approach for phenological attributes identification and prediction. To demonstrate its reliability, we followed a rigorous process in three main steps. Firstly, in Section 10.3.1, we describe the study area, the remote sensing data and the land cover classes used to assess our method. Then, in Section 10.3.2, we evaluate the goodness-of-fit of the model to non-smoothed MODIS spectral bands and VIs time series for 13 representative land cover classes in Portugal. In the course of this, we compare the goodness-of-fit of our model with that of two different standard approaches. Finally, in Section 10.3.3, using rice crops and shrublands as case studies, we show how to accurately estimate phenological key-markers from fitted responses and how to early predict important phenological attributes dates and magnitudes from our expected model adjusted to only few anterior observations.

10.3.1. Study area and experimental data set

Study area

We focused our experiences on the Portuguese mainland territory. The country has an area of approximately 89 000 km^2 and ground altitudes ranging from 0 to 2000 m above sea level. According to [23], in 2000 the agricultural areas occupied 37% of the country, closely followed by semi-natural areas (33%) and forest (27%), while artificial surfaces comprise only 3% of the total land cover. Portugal is in a transition region featuring diverse landscapes representing both Mediterranean and Atlantic climate environments. The study area is particularly well suited to evaluate data-model adequacy and to gauge the reliability of model-based approaches for identifying and predicting phenological attributes. Indeed, it encompasses natural and managed land cover types with different phenological characteristics, which are subject to different climatic, topographic and human induced forcing effects.

Earth observation data

Our study relies on the MOD09A1 product, an 8-day composite of seven surface reflectance images, freely available from MODIS Data Product website (http://modis.gsfc.nasa.gov). This specific product is an estimate of the surface spectral reflectance imaged at a nominal spatial resolution of 500 m as it would have been measured at ground level if there were no atmospheric scattering or absorption. The applied correction scheme compensates for the effects of gaseous and aerosol scattering and absorption, for adjacency effects caused by variation of land cover, for Bidirectional Reflectance Distribution Function (BRDF), for

coupling effects, and for contamination by thin cirrus [24]. We used a set of 46 MOD09A1 images for each of the seven spectral bands, covering a full year observation period, from January 2005 to December 2005. In addition, two VIs were also calculated for each date and used as additional band information, namely the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI).

As complementary information for reference sample land cover type's identification, we used orthorectified aerial images with four spectral bands in the blue, green, red and infrared wavelengths, and with 50 cm spatial resolution. These aerial images cover the whole Portuguese territory and were acquired during years 2004, 2005 and 2006.

Land cover classes

The thirteen land cover classes (Table 10.3) used in this study were derived from LANDEO nomenclature. This nomenclature was recently recommended by The Remote Sensing Unit (RSU) of the Portuguese Geographic Institute (IGP) to be used in regular production of user-driven land cover products for multi-scale environmental monitoring in Portugal [25]. The classes used in this study correspond mainly to vegetated land cover types because we want to demonstrate the relevance of our model-based approach for the identification and prediction of phenological attributes. Nevertheless, to illustrate the flexibility of our model, we also included land cover classes without intra-annual seasonality, namely Water (7) and Continuous artificial areas (11).

For each of the land cover classes presented in Table 10.3, we deterministically collected a reference sample of time series observations. Each observation corresponds to a MODIS pixel with 500 m nominal spatial resolution. Following the lines of [26], we overlaid a co-registered 500 m fishnet corresponding to the MODIS data grid with the aerial images for visual interpretation of reference land cover observations. To guarantee land cover homogeneity within each sample, observations for a given class were selected preferentially in geographic spots located in the centre of 3x3 MODIS pixels of homogeneous land cover. The goal was to reduce intra-class observations dissimilarities due to possible geometric disparities between multitemporal MODIS images. Nevertheless, we endeavored to spread, as much as possible, the observations over the mainland territory, in order to account for possible regional specificities of the classes, and to avoid geographic correlation between adjacent pixels [27].

By analogy with the medical imagery domain, we call voxel the time evolution of a measurement within a given pixel. In our case, we consider p=9 simultaneous measurements $\mathbf{x}^{(j)}$, j=1,...,p, corresponding to the reflectances in seven spectral bands and to two VIs. We

designate by $\mathbf{x}^{(j)} = \{x^{(j)}(t_i) : \ i = 1,...,N;\ 1 \leq t_i \leq T\}$ the *j*-th voxel and we group all the observations related to a pixel in a *p*-dimension vectorial time series noted $\mathbf{x}={}^{t}(\mathbf{x}^{(1)}, \mathbf{x}^{(2)},...,\mathbf{x}^{(p)})$. As we use 8-day composite satellite images, we have N=46 measurement points uniformly spaced along a T=365 days period.

Table 10.3. Land cover classes, label and respective number of collected reference observations.

Label	**Land cover class**	**Sample size**
7	Water	30
11	Continuous artificial areas	43
21	Non-irrigated herbaceous crops	40
22	Irrigated herbaceous crops	33
23	Rice crops	37
242	Vineyards	37
2411	Olive trees	48
35	Pastures	34
3121	Agro-forestry areas	42
34	Shrubland	46
3111	Cork tree forest	40
3112	Eucalyptus forest	34
321	Needleleaved forest	42

10.3.2. Data model adequacy

The primary objective of this section is to demonstrate the response of our model-based approach to diverse spectral reflectances and VIs time series for land cover classes of different phenology. We therefore apply our algorithm to the set of seven spectral reflectance bands and two VIs time series for each of the land cover classes introduced in Table III, and evaluate its goodness-of-fit. To position our method with respect to the current state-of-the-art alternatives, we compare its performances to those of Asymmetric Gaussian (AG) and Double Logistic (DL) functions presented in Section 10.2.1.

To evaluate models performances, we use the *Normalized Root Mean Square Error* (*NRMSE*). This measure was previously used by [11] and [28] to evaluate the fits of different models to time series of satellite images and can be expressed as a percentage, where lower values indicate smaller residual variance and thus a better fit. The *NRMSE* for each sample individual at spectral band or VI *j* is defined as:

$$NRMSE^{(j)} = \frac{RMSE^{(j)}}{\max\left\{x^{(j)}(t_i)\right\} - \min\left\{x^{(j)}(t_i)\right\}}, \text{ for } j = 1,\ldots,p. \qquad (10.10)$$

The $RMSE^{(j)}$ is calculated as:

$$RMSE^{(j)} = \left(\frac{SSE^{(j)}}{N-u}\right)^{1/2} \qquad (10.11)$$

where $SSE^{(j)} = \sum_{i=1}^{N}\left(x^{(j)}(t_i) - s_{\hat{\theta}}^{(j)}(t_i)\right)^2$, and u is the number of unknown model parameters. To estimate $s_{\hat{\theta}}(t_i)$ we start fitting the available time observations $\mathbf{x} = \{x(t_k): k \neq i;\ t_k \neq t_i\}$ with a parameter model of expression ((10.4). This standard cross validation process, usually used in longitudinal data analyses, yields an estimate of the expected fitting error when the respective time observation $x(t_i)$ is discarded from the adjusting set. The *NRMSE* is normalized on the range of the observed sample values at each spectral reflectance band or VI, and adjusted on the residual degrees of freedom of each model to penalize the models with more unknown parameters. Let us recall here that for AG and DL, the number of degrees of freedom depends directly on the number of extrema existing in the time series to be modelled, whereas it is systematically constant with our non harmonic model.

To avoid biased error estimates due to atmospheric noise in the time series, we removed cloudy observations from our sample data set, as similar as in [5] and [11]. This way, the *NRMSE* becomes also an estimate of the fitting error for the dates t_i identified with clouds and that are never used in the optimization process. Cloudy observations were identified with the surface reflectance quality assessment layer included in MOD09A1 products (see http://modis.gsfc.nasa.gov). Alternatively, we could also calculate the *NRMSE* from the model fitted time series $s_{\hat{\theta}}(t_i)$ and a smoothed observed time series (e.g. [28], [29]) or an expertly corrected observed time series (e.g. [3], [11]). However, these corrections on the observed $x(t_i)$ time series were discarded because they may introduce new unknown errors and dampen important phenological features in the real time series [30].

Table 10.4 gives the sample average *NRMSE*s corresponding to the nonlinear, AG and DL models' fits to multispectral reflectance and VIs time series for each land cover class. As a general remark, we can state that AG and DL fits to multispectral reflectances and VIs time series are similar for all land cover classes, and that the nonlinear model fits generally overcome the AG and DL fits. Indeed, the overall average *NRMSE* for the nonlinear model (5.25%) is significantly better than the overall average *NRMSE*s for the AG (6.97%) and DL (6.59%) functions at the 95% confidence level.

Table 10.4. *NRMSE*s (%) corresponding to the nonlinear, AG and DL functions' fits to spectral bands and VIs time series for each land cover class.

		Spectral band/VI								
Land cover class	**Function**	**b01**	**b02**	**b03**	**b04**	**b05**	**b06**	**b07**	**NDVI**	**EVI**
11	Nonlinear	7.52	4.53	6.62	7.14	5.88	7.71	5.97	1.76	2.12
	AG	10.35*	6.08*	9.42*	10.67*	7.76*	10.23*	7.73*	2.66*	2.94*
	DL	9.79*	5.88*	8.81*	9.37*	7.51*	9.52*	7.42*	2.41*	2.86*
21	Nonlinear	7.25	5.71	5.37	5.61	6.18	7.86	8.02	3.63	5.46
	AG	8.41*	7.02*	6.58*	6.89*	8.73*	9.89*	9.29*	3.74	5.30
	DL	8.41*	6.67*	6.41*	6.76*	8.08*	9.54*	8.52	3.40	5.01
22	Nonlinear	10.29	8.36	7.15	7.73	8.41	9.61	9.45	4.49	7.45
	AG	16.02*	10.57*	10.84*	11.15*	11.20*	14.04*	15.07*	6.06*	10.02*
	DL	14.63*	10.44*	9.45*	10.02*	10.14*	12.87*	13.13*	5.65*	8.99*
23	Nonlinear	6.67	7.00	5.36	5.44	9.22	9.30	8.93	4.19	5.68
	AG	11.52*	9.46*	8.89*	9.38*	12.76*	15.35*	12.94*	5.72*	8.81*
	DL	10.03*	9.05*	7.69*	8.20*	11.54*	13.6*	12.08*	5.66*	7.63*
2411	Nonlinear	5.71	4.63	4.77	4.85	5.57	6.39	5.90	2.25	3.07
	AG	6.75*	5.68*	5.94*	5.94*	7.07*	7.86*	7.23*	2.73*	3.92*
	DL	6.65*	5.40*	5.67*	5.78*	6.72*	7.68*	7.11*	2.49	3.49
242	Nonlinear	7.08	5.21	5.17	6.11	6.58	7.46	7.40	2.22	3.08
	AG	9.33*	6.79*	7.01*	7.84*	8.22*	10.01*	9.84*	3.26*	4.52*
	DL	8.99*	6.49*	6.51*	7.46*	7.89*	9.51*	9.05*	2.73*	3.96*
3111	Nonlinear	4.64	4.63	3.8	4.39	6.19	6.72	5.18	2.20	2.94
	AG	5.81*	5.72*	4.82*	5.47*	7.54*	8.47*	6.33*	2.76*	3.99*
	DL	5.68*	5.55*	4.74*	5.38*	7.28*	8.06*	6.21*	2.54*	3.70*
3112	Nonlinear	3.94	4.59	3.24	3.73	5.25	7.02	3.95	2.49	3.75
	AG	5.03*	5.90*	4.23*	4.81*	6.81*	8.88*	5.15*	3.20*	4.92*
	DL	4.94*	5.67*	4.17*	4.77*	6.61*	8.72*	5.02*	3.01*	4.77*
3121	Nonlinear	5.31	4.72	4.35	4.71	6.02	6.67	6.25	2.22	3.07
	AG	6.38*	5.74*	5.32*	5.84*	7.50*	8.03*	7.33*	2.48	3.65*
	DL	6.34*	5.45*	5.18*	5.69*	7.18*	7.79*	7.11*	2.39	3.29
321	Nonlinear	2.77	3.82	2.73	3.16	4.46	5.26	2.95	2.09	2.97
	AG	3.68*	4.73*	3.48*	4.04*	5.54*	6.75*	3.73*	2.52	3.84*
	DL	3.70*	4.52*	3.39*	3.98*	5.47*	6.50*	3.69*	2.41	3.60*
34	Nonlinear	4.28	3.93	3.40	4.05	5.36	5.86	4.63	2.07	2.55
	AG	5.48*	5.20*	4.34*	5.15*	6.99*	7.77*	5.93*	2.74*	3.35*
	DL	5.37*	5.01*	4.29*	4.99*	6.74*	7.51*	5.72*	2.36*	3.15*
3512	Nonlinear	4.16	4.63	3.42	4.04	6.42	7.05	5.13	1.76	3.01
	AG	6.38*	6.27*	4.63*	5.21*	8.51*	8.44*	6.20*	2.76*	4.24*
	DL	5.02*	5.86*	4.24*	5.00*	8.41*	8.69*	6.76*	2.63*	4.42*
7	Nonlinear	5.70	3.67	5.81	5.51	5.00	9.45	3.48	7.47	2.64
	AG	7.78*	4.82*	7.70*	7.07*	7.02*	13.43*	4.83*	8.93*	3.64*
	DL	7.22*	5.54*	7.20*	6.67*	7.15*	12.49*	4.56*	8.39*	3.35*

* Significantly different at 5% from the correspondent nonlinear *NRMSE*.

More formally, a closer view at Table 10.4 shows that the average *NRMSE*s for VIs time series fits are better than the *NRMSE*s for multispectral reflectances time series fits. These results are mainly due to the fact that not all clouds are identified by the MODIS surface reflectance quality assessment layer, and some sharp and localized atmospheric disturbances still pollute the data classified as clear. Accordingly, because VIs are less sensitive than spectral reflectance bands to cloud contamination and also minimize many contamination problems present in the spectral wavelength bands, such as those associated with aerosol influences [31], the residuals of the smoothed curves are smaller for VIs.

Now, looking at land cover classes, we see that the nonlinear model fits significantly overcome the AG and DL fits for Water (7) and Continuous artificial areas (11) because these functions often fail at reproducing time series without apparent seasonal waveforms [5], [11]. Still, as AG and DL functions are not fully versatile and their use is generally constrained to VIs time series of land cover classes with seasonal waveforms [5], [9], [11], [17], [18], we expected those functions to perform as good as the nonlinear model for those datasets. Conversely, we note that the nonlinear model fits to VIs time series for agricultural cover types, such as Rice crops (23), are significantly superior to the respective DL and AG fits. From Figure 10.3, we remark that the AG and DL fits describe the EVI data as good as the nonlinear fits in the time interval around the maximum. At the limbs, however, the fits are less good than the nonlinear fits. The reason is that because AG and DL functions are piecewise, they cannot reproduce the global seasonal pattern of the class from a single yearly time series set of observations. We remark that neither AG nor DL functions are able to depict properly the local maximum (respectively local minimum) in the left half (respectively right half) of the observed curves. This drawback was presented before by [5], which propose the use of three consecutive years of data to correctly model the limbs of the seasons for the central year. However, if we do not have data for the surrounding years, this problem makes difficult the use of AG and DL functions for fitting satellite images time series.

Importantly also, Figure 10.3 shows that the nonlinear model fits present identical time patterns for all individuals of the same land cover class, in opposition to AG and DL fits. We verify this also for natural vegetated land cover classes. For example, AG and DL fits to EVI time series for Needleleaved forest (321) individuals are considerably heterogeneous (Fig. 4). On the other hand, the nonlinear fits to the same individuals have consistent time trajectories and fits are independent of data disturbances not depicted by the quality assessment layer. This outcome is extremely important for phenological attributes

identification and land cover classes' characterization through time series analysis of satellite images.

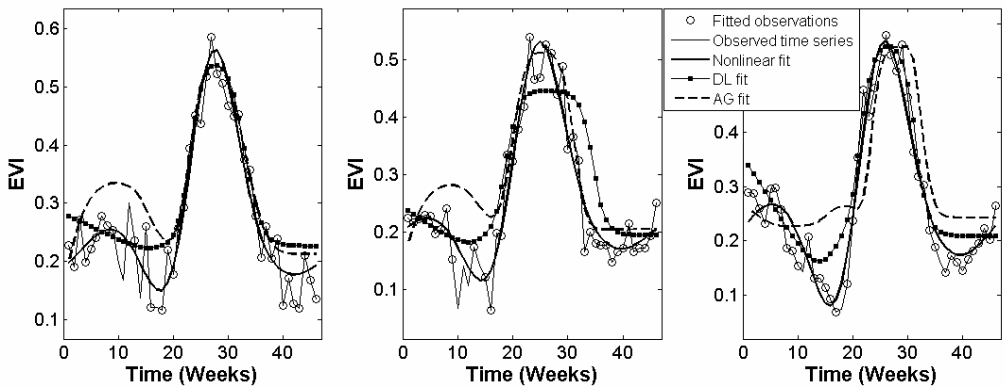

Figure 10.3. Observed EVI time series for three Rice crops (23) individuals and the respective nonlinear, AG and DL fits.

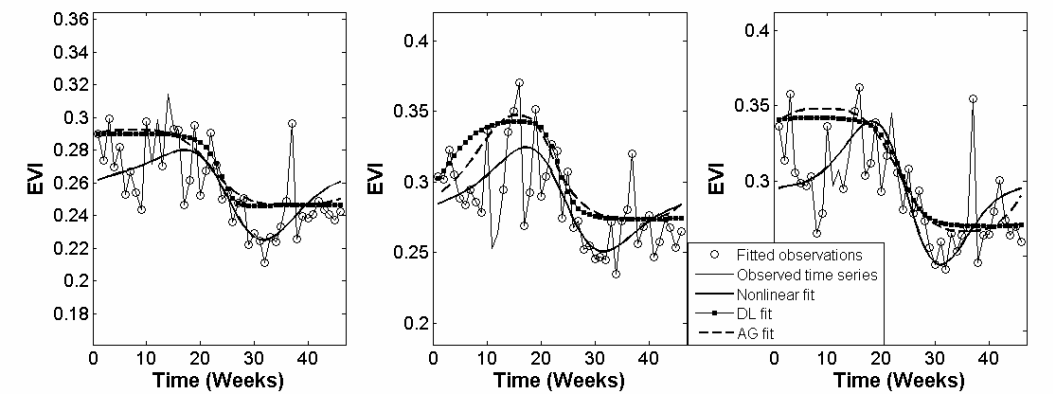

Figure 10.4. Observed EVI time series for three Needleleaved forest (321) individuals and the respective nonlinear, AG and DL fits.

We also perceive that the border values of the annual curves fitted by AG and DL functions do not match (Figure 10.4). We expected the waveforms of natural land cover classes to present similar EVI values at the beginning and ending times of the year because we are dealing with individuals that have recurring seasonal patterns across different years – unless some unexpected environmental forcing effect occurs during the modelled year, which is never the case here. However, because AG and DL functions are piecewise, they cannot reproduce reliable fits using a single yearly time series set of observations.

The nonlinear model fits to VIs time series for most natural vegetated cover types with slow time variations, also present significantly lower fitting residuals. This is the case of, e.g. models fits to EVI time series for Eucalyptus forest (3112) observations (Figure 10.5). As pointed by [11], the reason is that AG and DL functions often fail at reproducing time series that do not correspond to the shape of the Gaussian or Logistic curves.

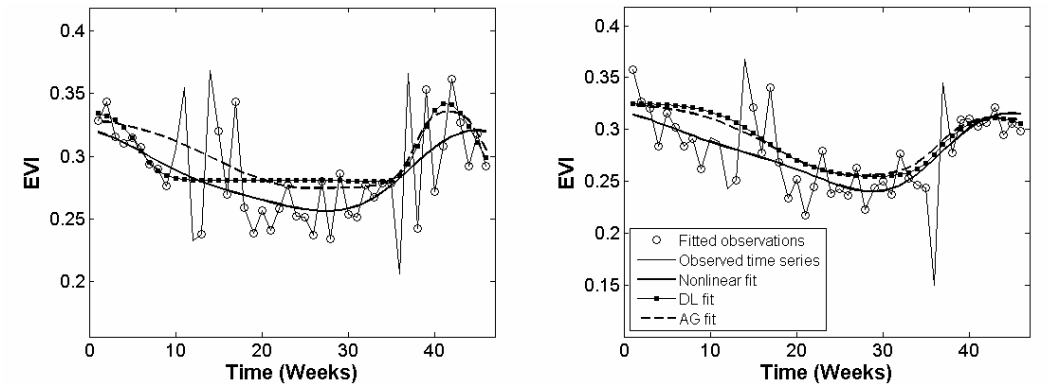

Figure 10.5. Observed EVI time series for two Eucalyptus forest (3112) individuals and the respective nonlinear, AG and DL fits.

10.3.3. Phenological key-markers identification and prediction

Identification

Now that we confirmed the ability of the nonlinear model to fit multispectral reflectances and VIs time series for diverse land cover classes, we will use the modelled curves to effectively identify phenological measures. This is a common approach and several methods for deriving annual phenological metrics from smoothed time series have been proposed in the literature, e.g. [9], [32]-[35]. These methods aim at overcoming data noise corruptions and at providing a more accurate continuous time evolution of land cover phenological characteristics that are necessarily discrete due to sampled satellite image measurements [4].

Following the lines of [9], we use the first and second derivatives of our modelled VIs time series to identify the inflection points of the intra-annual curves and, accordingly, the unbiased phenological attributes estimates. The number and position of the inflection points depend on the seasonal characteristics of the land cover class being modelled and on the respective physical measure used to derive the phenological attributes.

To illustrate our approach, we use the Rice crops (23) NDVI time series observations from our sample database. For example, the minimum of the second derivative of the modelled time series is the time for maximum greenness (i.e. heading) t^M of the crop, while the

maximum (respectively the minimum) of the first derivative corresponds to the onset t^{OnM} (respectively to the offset t^{OfM}) time of maturity. In the sequence, we identify the times for transplant t^{T} and harvesting t^{H} as the second derivative maxima. An identification example of phenological attributes using a Rice crops (23) NDVI time series sample observation is shown in Figure 10.6.

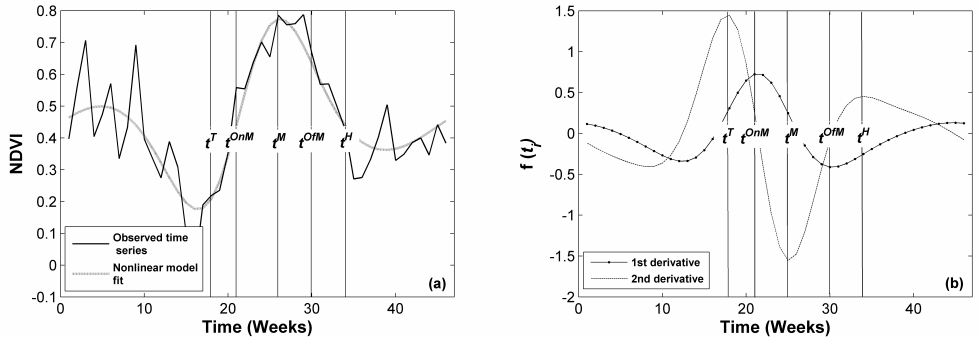

Figure 10.6. Rice crops (23) phenological attributes. a) NDVI time series and respective model fit for a single observation; b) first and second derivatives of modelled time series as function of time. Vertical bars represent the times for: Transplanting t^{T}; Onset of maturity t^{OnM}; Maximum greenness t^{M}; Offset of maturity t^{OfM}; Harvesting t^{H}.

To evaluate the accuracy of our analytical method, we compared the model-based estimated phenological times for the Rice crops (23) observations of our sample database to the empirical reference dates provided in Portuguese Agricultural Calendar [36]. More formally, we followed a rigorous process in two main steps: 1) we fitted each NDVI time series sample observation (corresponding to geographically distant pixels) with our smooth model; 2) we estimated three specific phenological key-dates for each sample observation through the inflection points of the first and second derivatives of each continuous fitted curve. The estimated key-dates, expressed in week of the year (WOY) are: the rice crops transplant date t^{T} (between the 15th and the 19th WOY), the maximum greenness time t^{M} (between the 24th and the 28th WOY), and the harvesting date t^{H} (between the 33rd and the 37th WOY). Figure 10.7 shows the estimated Rice crops (23) sample probability distributions for the three phenological dates. In all the cases, not only the underlying phenological measures are reasonably estimated, because the mean dates correspond to the theoretical references of [36], but also the sample estimated phenological key-dates are normally distributed and have very small standard deviations from the respective expected date.

Prediction

Suppose now that we want to predict characteristic phenological attributes from only a series of anterior images of the same year. Let us call $\mathbf{x} = \{x(t_i)\colon i = 1,..., n < N;\ 1 \le t_i < T\}$ the observed intra-annual time series corresponding to the evolution of a reflectance or of a VI in a given pixel of class C_l. To predict the date of interest t_m, with $t_n < t_m \le t_N$, we start fitting the available observations **x** with a parameter model of expression (10.4). To cope with censured data $x(t_k)$: $t_k > t_n$ that may mislead the initial guess $\hat{\theta}_0$, we initialize the solving procedure of Section 10.2.2 - **Parameters identification**, with an average meta-parameter $\bar{\theta}_l$ representative of the class C_l (the class model parameters are empirically determined from a training sample set). Once estimated the fitting function $s_{\hat{\theta}}(t)$, its evaluation for $t > t_n$ serves as a prediction for the phenological development of the pixel area, and forecasted milestone dates associated to this particular land cover class simply derive from the variation analysis of the path $s_{\hat{\theta}}(t)\colon t > t_n$.

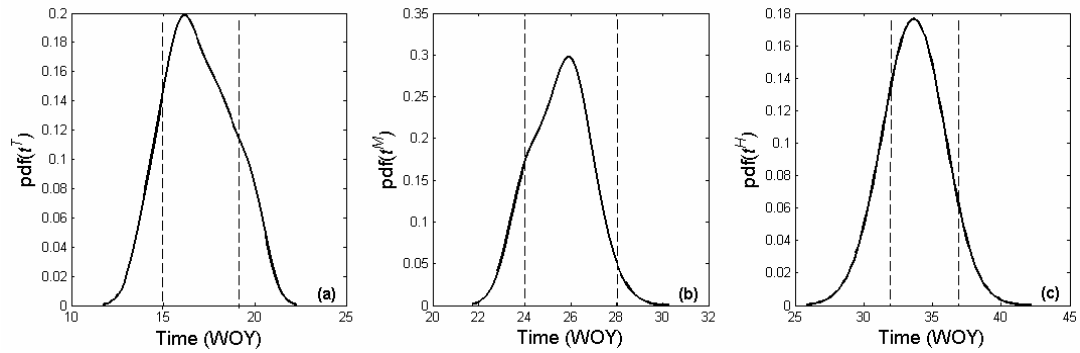

Figure 10.7. Estimated Rice crops (23) sample probability distributions for: a) Transplanting time t^T; b) Maturity time t^M; c) Harvesting time t^H. Vertical bars represent the limits for the phenological reference key-dates provided in [36] for rice crops in Portugal.

Let us stress that this approach holds only with global models that uniformly fit the time series over a full year period. Analytic models of the form of (10.4) naturally meet this requirement: structurally constrained by definition, they are consistently identifiable from a time incomplete set of observations, while sufficiently versatile to yield patterns able to match a broad spectrum of land cover classes and physical measurements. That is not the case though with models such as Asymmetric Gaussian (AG) or Double Logistic (DL) functions whose parameters are piecewise constant and hence unpredictable from one segment to the next.

We illustrate the relevance of our model-based predictive approach with two significant case studies. The first application deals with rice crops phenological attributes prediction. Reference [29] state that essential rice crop's phenological stages for evaluating crop productivity and crop management are the planting, the heading and the harvesting dates. Indeed, vegetation magnitudes at the heading are related to potential yield provided that no accident occurs after that stage [4]. As it seems unrealistic to predict the planting date from phenological arguments only (it is mainly determined by human decisions), our analyses focus on the prediction of heading and harvesting dates. The second application concerns the prediction of shrubland phenological attributes. Our goal here is to predict both the time and the magnitude of vegetation minimum greenness before the Portuguese fire season, which officially starts in June (20^{th} WOY, [37]). Indeed, because fire is sensitive to the quantity of green vegetation, prescribed burning activities can benefit from the assessment of the minimum vegetation greenness extent and timing [38]. Therefore, if these phenological attributes are forecasted before fire season, then it will become easier to plan fire-fighting activities and to allocate surveillance infrastructures in due time. The main difference in these prediction case studies is related to political constrains. On the one hand, agricultural monitoring undergoes no political constraints, and prediction can be repeated at any time during the crop growing season in order to progressively refine phenological attributes estimates. On the other hand, planning of an efficient forest fire prevention policy imposes an official date *t* for predicting important phenological attributes of natural vegetation.

Rice crops (23) case study: In Section 10.3.3 – **Identification**, we considered a set of NDVI time series sample observations representative of rice crops, and for each of these we estimated the phenological dates of maturity t^M and harvesting t^H. We now use these estimates as the target values that we want to predict.

We remove from our dataset all observations that lie beyond the current date t_n, and we apply our prediction procedure to all time series censured in this way. We note $t_n^{\hat{M}}$ (respectively $t_n^{\hat{H}}$) the predicted values to make their conditional dependence on observations up to date t_n explicit. The prediction error $eT_n^M = t^M - t_n^{\hat{M}}$ (respectively $eT_n^H = t^H - t_n^{\hat{H}}$) is a random variable whose statistics (mean, variance...) assess the relevance of our model through the quality of its prediction.

In Figure 10.8, we present the *Root Mean Square Prediction Errors* (*RMSPE*) for the dates of maturity t^M and for the dates of harvesting t^H as functions of t_n. *RMSPE*s were computed

for all Rice crops (23) sample individuals and for $t_n = t^M - k, k = 1,...,10$ weeks (respectively $t_n = t^H - k, k = 1,...,10$ weeks).

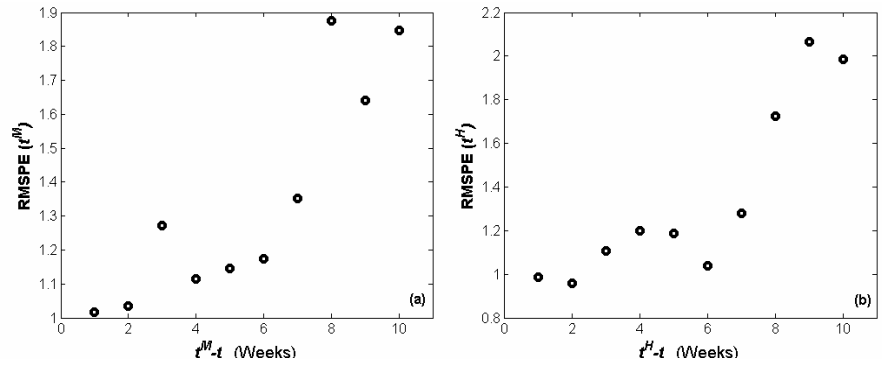

Figure 10.8. Rice crops (23) sample *RMSPE*s for: a) dates of maturity t^M as function of t_n; b) dates of harvesting t^H as function of t_n.

From our experiments, we notice that the prediction errors for both t^M and t^H estimates decrease as k decreases, i.e. as $t_n \rightarrow t^M$ (respectively $t_n \rightarrow t^H$). More importantly though, with six weeks of antecedence (i.e. k = 6), we are able to predict the dates of maturity t^M with ± 1.1 weeks of precision; in the same way, the *RMSPE* for forecasted dates of harvesting t^H is of approximately ± 1.3 weeks when forecasting seven weeks ahead (i.e. k = 7).

Assume now that $NDVI^M$ denotes the maximum greenness for Rice crops (23) sample observations at the estimated phenological dates of maturity t^M, and $NDVI_n^{\hat{M}}$ refers to the forecasted maximum greenness values based on a model fitted on observations up to date t_n. Considering k = 6 for all observations, we compute the sample estimated probability distribution for the maximum greenness prediction error $eN_n^M = NDVI^M - NDVI_n^{\hat{M}}$, and present the results in Figure 10.9.

From Figure 10.9, we see that the estimated sample probability distribution for the maximum greenness prediction errors is normal and approximately zero mean, with a standard deviation of 0.05. Moreover, considering that the average sample $\overline{NDVI^M}$ is 0.79, then the estimated standard deviation is only 6% that value. This result confirms that our model-based prediction approach provides 6 weeks ahead precise crop development information for time-specific crop management.

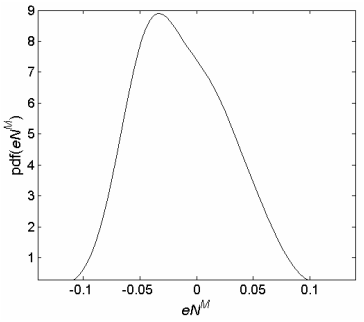

Figure 10.9. Estimated Rice crops (23) sample probability distribution for eN_n^M, where $t_n = t^M - 6$.

Shrublands (34) case study: Now, consider that we want to predict the minimum greenness value $NDVI_n^{\hat{m}}$ for Shrubland (34) observations and its respective date $t_n^{\hat{m}}$ of occurrence based on a model fitted only to tn past observations of the NDVI time series. In our case, the prediction process is limited to t_n = 20 past observations to respect the time constrain on the forest fire prevention program in Portugal. Following the same strategy as for rice crops, we estimate the reference phenological dates of minimum greenness t^m and the respective minimum greenness $NDVI^m$ for Shrubland (34) sample observations using our analytical method based on the first and second derivatives of the modelled time series. In the sequence, we calculate the estimated sample probability distributions for the prediction errors $eT_{20}^m = t^m - t_{20}^{\hat{m}}$ and $eN_{20}^m = NDVI^m - NDVI_{20}^{\hat{m}}$ (Figure 10.10).

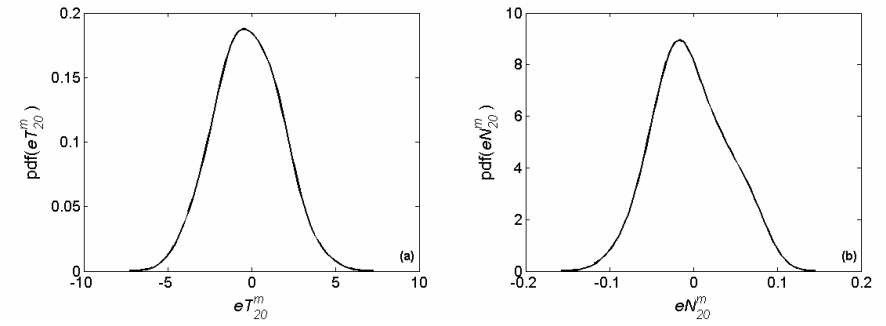

Figure 10.10. Estimated Shrubland (34) sample probability distributions for: a) eT_{20}^m; b) eN_{20}^m.

The minimum greenness time prediction errors eT_{20}^m for Shrubland (34) sample observations have an estimated zero mean normal probability distribution with a standard deviation of ±2 weeks. Considering that the sample average $\overline{t^m}$ is 10 weeks after the prediction date t_{20}, we

can state that the estimated standard deviation for the prediction error is small and that the prediction is accurate. Moreover, the minimum greenness prediction errors eN_{20}^{m} have an estimated zero mean normal probability distribution with a standard deviation of 0.03. We remark that this value is only 9% the observed average $\overline{NDVI^{m}}$, thus confirming that $NDVI_{20}^{\hat{m}}$ is precise and can effectively be used for fire-fighting planning strategies.

In Figure 10.11, we present the *NDVI* time series model fits for two Shrubland (34) observations and the respective predicted time series $NDVI_{\theta}(t)$, $t > t_{20}$. The visual analysis of forecasted time series suggest that the model-based predictions do not mislead the planning actions to be performed for Shrubland (34) observations with divergent greenness magnitudes.

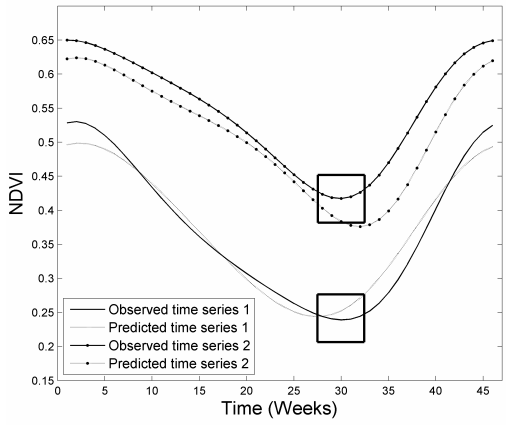

Figure 10.11. *NDVI* time series model fits for two Shrubland (34) individuals and respective model predictions for t_{20}. The rectangles represent the standard deviations for eT_{20}^{m} and eN_{20}^{m}.

10.4. Conclusion

In this paper we proposed and evaluated a model-based approach for the identification and prediction of vegetation phenological attributes. The method relies on a consistent, whereas very simple, nonlinear parametric harmonic model that we developed in a related work [16] to fit intra-annual multispectral reflectances and VIs time series from satellite images. The identifiability study revealed that the model parameters are easily and perfectly estimated from a single yearly time series set of observations, and that the sensitivity of the estimation procedure is independent of parameters magnitude. We demonstrated that the model parameters are reliably identified from continuous signals sampled on different time lattices,

as long as the intra-annual time variations of the underlying signals are captured by the quantization step. Moreover, the noise power of model fits to time series synthesized with different SNR values is always lower than the noise power of the respective input data, thus indicating that the model is able to diminish disturbances intensity in time series.

The goodness-of-fit confirmed that the nonlinear harmonic model is sufficiently able to match multispectral reflectances and VIs time series derived from satellite images for disparate land cover classes. Moreover, the nonlinear harmonic model can be used on all time series data sets without supervision and prior testing, because it does not need to be tuned for specific land cover classes and remotely-sensed measurements. From a general viewpoint, the nonlinear harmonic model confirmed to be more versatile than AG and DL functions to yield a broad spectrum of physical measurements describing the intra-annual time evolution of land surface characteristics. In particular, the nonlinear harmonic model performed even better than the other fitting algorithms for some unexpected situations, such as for modelling VIs time series of vegetated land cover classes, which are the focus of AG and DL functions. The nonlinear harmonic method, as outlined in this paper, holds the additional advantage that the time series is fitted over a full year period. Piecewise functions, such as AG and DL, fail at accurately matching the global waveform of numerous time series because they rely on the linear combination of local and independent intra-annual functions.

The original model-based applications presented here for the identification and prediction of phenological key-attributes revealed extremely successful under particularly difficult experimental conditions. The identification of phenological key-markers using first and second derivatives of the nonlinear harmonic model proved reliable and functional for the study cases. From our experiments, the estimated key-attributes perfectly match the respective reference values for the study area. More important though, the prediction approach revealed that the model is consistently identifiable from a time incomplete set of observations and flexible enough to allow for forecasts of different phenological attributes for a broad spectrum of land cover classes. This particularly interesting and promising application allowed for the prediction up to ten weeks ahead of important phenological attributes for agricultural management and pre-season planning of forest fire fighting activities and surveillance from only few anterior observations.

Furthermore, we remark that the model can be applied in numerous different applications. Because the five model parameters contain information about the continuous intra-annual variation of cover types' reflectances, we can use the model for reducing the dimensionality

of satellite images time series and the model parameters for multitemporal land cover classification.

10.5. Acknowledgment

Research by Hugo Carrão was founded by the "Fundação para a Ciência e Tecnologia" (SFRH/BD/18447/2004). Research by Paulo Gonçalves and Mario Caetano was partially supported by the "Luso-French Integrated University Actions – 2007-2008". This work was performed while Hugo Carrão was visiting INRIA-ENS-Lyon, LIP (RESO), France.

10.6. References

[1] P. J. Pinter *et al.*, "Remote Sensing for Crop Management," *Photogramm. Eng. Remote Sens.*, vol. 69, pp. 647-664, 2003.

[2] B.A. Bradley, R.W. Jacob, J.F. Hermance and J.F. Mustard, "A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite NDVI data," *Remote Sens. Environ.*, vol. 106, pp. 137–145, 2007.

[3] J. F. Hermance, R. W. Jacob, B. A. Bradley and J. F. Mustard, "Extracting Phenological Signals from Multi-Year AVHRR NDVI Time Series: Framework for Applying High-Order Annual Splines with Roughness Damping," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, pp. 3264-3276, October 2007.

[4] S. Moulin, A. Dondeau and R. Delecolle, "Combining agricultural crop models and satellite observations from field to regional scales," *Int. J. Remote Sens.*, vol. 19, pp. 1021–1036, 1998.

[5] P. Jönsson and L. Eklundh, "Seasonality extraction by function fitting to time-series of satellite sensor data," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, pp. 1824–1832, August 2002.

[6] H. Carrão, P. Gonçalves and M. Caetano, "Contribution of multispectral and multitemporal information from MODIS images to land cover classification," *Remote Sens. Environ.*, vol. 112, pp. 986-997, 2008.

[7] Moody and D. M. Johnson, "Land-surface phenologies from AVHRR using the discrete Fourier transform," Remote Sens. Environ., vol. 75, pp. 305–323, 2001.

[8] Z. T. Li and M. Kafatos, "Interannual variability of vegetation in the United States and its relation to El Nino/Southern Oscillation," *Remote Sens. Environ.*, vol. 71, pp. 239–247, 2000.

[9] X. Y. Zhang *et al.*, "Monitoring vegetation phenology using MODIS," *Remote Sens. Environ*., vol. 84, pp. 471–475, 2003.

[10] K.R. McCloy and W. Lucht, "Comparative evaluation of seasonal patterns in long time series of satellite image data and simulations of a global vegetation model," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, pp. 140-153, January 2004.

[11] P. S. A. Beck, C. Atzberger, K. A. Høgda, B. Johansen and A. K. Skidmore, "Improved monitoring of vegetation dynamics at very high latitudes: a new method using MODIS NDVI," *Remote Sens. Environ.*, vol. 100, pp. 321-334, 2006.

[12] R. E. Burgan *et al.*, "WFAS: wildland fire assessment system," *Fire Management Notes*, vol. 57, pp. 14-17, 1997.

[13] M. Caetano, S. Freire and H. Carrão, "Fire risk mapping by integration of dynamic and structural variables," in *Remote Sensing in Transition*, R. Goossens, Ed. Roterdam: Millpress, 2004, pp. 319-326.

[14] J. Meyer-Roux and C. King, "European achievements in remote sensing: agriculture and forestry," *Int. J. Remote Sens.*, vol. 13, pp. 1329-1341, 1992.

[15] C. L. Wiegand *et al.*, "Development of Agrometeorological Crop Model Inputs from Remotely Sensed Information," *IEEE Trans. Geosci. Remote Sens.*, vol. 24, pp. 90-98, January 1986.

[16] P. Gonçalves, H. Carrão and M. Caetano, "Parametric model for intra-annual reflectance time series," in *New Developments and Challenges in Remote Sensing*, Z. Bochenek, Ed. Roterdam: Millpress, 2007, pp. 517-526.

[17] P. Jönsson and L. Eklundh, "TIMESAT — A program for analyzing time-series of satellite sensor data," *Comput. Geosci*., vol. 30, pp. 833–84, 2004.

[18] J. I. Fisher, J. F. Mustard and M. A. Vadeboncoeur, "Green leaf phenology at Landsat resolution: Scaling from the field to the satellite," *Remote Sens. Environ.*, vol. 100, pp. 265−279, 2006.

[19] M. Barnsley, *Fractals Everywhere*, NewYork: Academic Press Inc., 1988.

[20] K. Levenberg, "A Method for the Solution of Certain Problems in Least Squares," *Quart. Appl. Math*., vol. 2, pp 164-168, 1944.

[21] D. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *J. Appl. Math*., vol. 11, pp. 431-441, 1963.

[22] K. Formann, "Linear logistic latent class analysis for polytomous data," J. Amer. Statistical Assoc., vol. 87, pp. 476–486, 1992.

[23] M. Caetano, H. Carrão and M. Painho, *Alterações da ocupação do solo em Portugal Continental: 1985 – 2000*, Amadora, Portugal: Instituto do Ambiente, 2005.

[24] E. F. Vermote and A. Vermeulen. (2005, October 17), *Atmospheric correction algorithm: Spectral reflectances (MOD09) algorithm technical background document* [Online]. Available: http://modis.gsfc.nasa.gov/data/atbd/.

[25] H. Carrão, P. Gonçalves and M. Caetano, "Multitemporal MERIS images for land cover mapping at national scale: the case study of Portugal," *Int. J. Remote Sens.*, to be published.

[26] R. S. Defries, M. Hansen, J. R. G. Townshend and R. Sholberg, "Global land cover classifications at 8km spatial resolution: The use of training data derived from Landsat imagery in decision tree classifiers," *Int. J. Remote Sens*., vol. 19, pp. 3141-3168, 1998.

[27] T. O. Hammond and D. L. Verbyla, "Optimistic bias in classification accuracy assessment," *Int. J. Remote Sens.*, vol. 17, pp. 1261-1266, 1996.

[28] C. Bacour, F.M. Bréon and F. Maignan, "Normalization of the directional effects in NOAA-AVHRR reflectance measurements for an improved monitoring of vegetation cycles," *Remote Sens. Environ.*, vol. 102, pp. 402-413, 2006.

[29] T. Sakamoto *et al.*, "A crop phenology detection method using time-series MODIS data," *Remote Sens. Environ*., vol. 96, pp. 366–374, 2005.

[30] J. I. Fisher and J. F. Mustard, "Cross-scalar satellite phenology from ground, Landsat, and MODIS data," *Remote Sens. Environ.*, vol. 109, pp. 261−273, 2007.

[31] N. Pettorelli *et al.*, "Using the satellite-derived NDVI to assess ecological responses to environmental change," *Trends Ecol. Evol*., vol. 20, pp. 503–510, 2005.

[32] B. C. Reed *et al*., "Measuring phenological variability from satellite imagery," J. Veg. Sci., vol. 5, pp. 703-714, 1994.

[33] S. Moulin, L. Kergoat, N. Viovy and G. Dedieu, "Global-Scale Assessment of Vegetation Phenology Using NOAA/AVHRR Satellite Measurements," J. Climate, vol. 10, pp. 1154–1170, June 1997.

[34] M. A. White, P. E. Thornton and S.W. Running, "A continental phenology model for monitoring vegetation response to interannual climatic variability," Global Biogeochem. Cycles, vol. 11, pp. 217–234, 1997.

[35] B. Duchemin, J. Goubier and G. Courrier, "Monitoring phenological key-stages and cycle duration of temperate deciduous forest ecosystems with NOAA-AVHRR data," Remote Sens. Environ., vol. 67, pp. 51–67, 1999.

[36] M. F. Ripado, *Calendário Rural*, Lisboa, Portugal: Litexa Editora, 1991.

[37] DGRF – Direcção Geral dos Recursos Florestais. (2005, June 28), *Inventário Florestal Nacional* [Online]. Available: http://www.dgrf.min-agricultura.pt/ifn/index.html.

[38] R. E. Burgan and R. A. Hartford, "Monitoring vegetation greenness with satellite data," USDA Forest Service, Intermountain Research Station, Ogden, UT, Gen. Tech. Rep. INT-297, 1993.

11 CONCLUSIONS

The technologies and methods of remote sensing have evolved dramatically to include a suite of sensors operating at a wide range of imaging scales and time frequencies with potential interest and importance to planners and land managers. Coupled with the ready availability of historical remote sensing data, the reduction in data cost and increased resolution from satellite platforms, remote sensing technology appears poised to make an even greater impact on planning agencies and land management initiatives involved in characterizing and monitoring land cover at a variety of spatial scales. Clearly, however, the fast-paced developmental nature of remote sensing technology often overlooks the needs of end-users as it continues to outpace the accumulation of experience and understanding. As a result, effective real-world operational examples of land cover characterization and monitoring from satellite images remain relatively rare.

In this research we aimed at giving response to some of the land cover characterization problems endorsed to the Portuguese Geographic Institute (IGP) in the framework of the national land cover mapping program launched in 2004 by the Portuguese R&D National Strategy for Space, namely the planning of strategies for the regular mapping of land cover distribution and analysis of land cover dynamics at the national scale. The goal of this investigation was aimed primarily at developing Earth Observation (EO)-based products responding to real IGP needs. The definition, implementation and validation of these products was carried out in close collaboration with this end-user organization, and in agreement with its standards and practices. To tackle the land cover characterization problems, we focused on the study of multispectral time series data sets of medium spatial resolution satellite images covering the whole Portuguese territory. We followed a rigorous process in three main steps: 1) assessment of the multitemporal information contribution for land cover classes discrimination at the national scale; 2) definition of an exhaustive land cover mapping protocol for the automatic production of a land cover map with a set of multitemporal satellite images and respective accuracy assessment; and 3) development of a model to fit time series sets of multispectral satellite images for optimizing time information usage for the characterization of land cover distribution and land cover monitoring.

11.1. Main outcomes

From a general viewpoint, the analysis of the contribution of multitemporal satellite images information for land cover classes' discrimination (Chapter 2) revealed that multispectral data acquired at indiscriminate intra-annual times does not significantly improve the overall

classification accuracy from single date multispectral information. The small contribution of multispectral information acquired at different times for the overall classification accuracy is mainly due to the fact that only the discrimination between a small number of cover types benefit from the use of multiple input time features. Nevertheless, the combinatorial use of multispectral information from multiple dates, namely one summer and one winter image, demonstrated to be important for the discrimination of specific land cover classes. Therefore, as long as the classifier supports the features' dimension of the input data set, the classification results gain with the use of multitemporal data.

In the sequence, we confirmed that the spectral information from satellite images is more important than the respective temporal information for broad land cover classes' discrimination in the study area (Chapter 3). In fact, from the viewpoint of the overall accuracy, as long as the whole multispectral information for the most discriminant yearly time (i.e. summer in the case of our study area) is used as input data for the classification task, the multitemporal information can be disregarded. More importantly though, the results indicate that the overall land cover classification accuracy attained with an intra-annual time series set of a single spectral measure is always significantly lower than the accuracy achieved with the complete multispectral data set for a single date. This outcome suggests that even resorting to yearly time series sets of satellite images, it is impossible to discriminate between the cover types of interest if the appropriate spectral information is not available. Indeed, we verified that it is necessary to use a complete yearly time series set of at least three spectral measures to match the classification accuracy attained with the complete multispectral data set for a single date.

From a user's perspective, these results are extremely important. The fact that we do not need all of the yearly images composites to effectively improve the discrimination between cover types, facilitates the land cover classification tasks. On the one hand, the use of fewer input features for classification makes possible the use of standard classifiers that do neither need to be tuned by end users and are available in most widespread Earth Obervation (EO) data analysis software. On the other hand, satellite images processing times are reduced, as well as the associated images acquisition, storage and correction costs. Moreover, the size of the training sample that is required to deal with a smaller input dimensional feature space, in the context of a supervised classification, is also reduced. The reduction in the training sample size corresponds to a reduction in the sample acquisition time and cost.

Nonetheless, for other types of EO-based applications, e.g. forest fires prevention and crop yields estimation, there is the need to use a complete intra-annual time series set of satellite

images as input features. In Chapter 10, we focused on the development of a nonlinear harmonic model for fitting multispectral and vegetation indices time series from satellite images. The goal was to enhance the contribution of multitemporal information from satellite images for the land cover characterization, while reducing the subjective involvement of non-expert end users. From an operational viewpoint, the model is consistent, very simple, parsimonious, readily identifiable, robust to noise and versatile. These properties make it advantageous over other standard modeling approaches. The continuous multispectral time trajectories of land cover individuals can be summarized into five model parameters that contain the key time discriminant information from the whole original discrete time series datasets. Because the number of model parameters is parsimonious, the model outputs are effortlessly handled from a computational viewpoint and may be straightforwardly used in the framework of classification tasks performed with standard classifiers. Moreover, because our nonlinear harmonic model is global and uniformly fits the time series over a full year period, it is consistently identifiable from a time incomplete set of observations. The experimental results suggest that the model can be used as an approach to derive phenological attributes in early time, to serve as input features for decision systems, and to support policy makers within their decision-making frameworks. We can use the model-based prediction approach for the effective estimation, in early time, of phenological features: to derive crop yields estimates; to serve as input for agro-climatic models; and to detect abnormal crop development cases.

We suggested also that the use of standard smoothing techniques for the correction of local disturbances in satellite images time series must be avoided or at least specifically planned for each application (Chapter 10). The problem is that it is generically impossible to know the true signal magnitude under the noise shot, unless for some pixels' areas that are also monitored at the surface. Thus, local corrections on the time series, performed with standard smoothing techniques, may introduce new unknown errors and even augment the noise magnitude. This is always the case in denoising problems. Yet, people keep working on new techniques for it is a very challenging issue. In this dissertation, we invoked two solutions to deal with this problem: 1) to use a model that make a global fit of the time series, such as our nonlinear harmonic model; 2) or simply to remove the affected data from the time series based on satellite images quality assessment layers. The last solution is not as good as the first because it will reduce both the size of the data set available for land cover characterization and the range of analyses that can be performed from an incomplete and irregularly spaced time series data set. Thus, we promote the use of the nonlinear harmonic model to reduce the intensity of disturbances in the time series. As the model converts the

discrete time series measurements into a continuous measurement that is a simple function of time, it will in addition ease the estimation of the physical measures magnitudes for dates that were removed or not recorded by the satellite sensor.

Other important contributions of this research concern the approach for the automatic production of a land cover map for Continental Portugal from medium spatial resolution satellite images and the respective accuracy assessment strategy (Chapter 9). Specifically, we developed a land cover map with 16 land cover classes for the study area with an overall accuracy of 80% ± 2% at the 95% confidence level. Moreover, the estimated user's accuracies were, for almost all classes, superior to 60% and the respective absolute precisions estimates around 10% at the 95% confidence level. These results suggest that medium spatial resolution satellite images can be effectively used for the automatic land cover mapping at the national scale. Indeed, the thematic accuracies estimated for our map are consistent, after nomenclatures synchronization, with the accuracies of land cover maps produced by visual interpretation of high spatial resolution satellite images for the same study area, such as the CLC2000. The success of our work is due to the result of numerous independent investigations performed in the previous Chapters, namely: the selection of the most adequate set of EO data (Chapter 4); the classification strategy (Chapters 5 and 6); and the definition of an adequate sampling protocol for accuracy estimators computation and testing sample collection (Chapter 7 and 8).

Regarding EO data, we used a time series set of satellite images acquired by MERIS sensor for the year 2005. In Chapter 4, we confirmed that MERIS data are more suitable than MODIS data for discriminating between the land cover classes that best characterize the Portuguese landscape. The results indicate that the differences in land cover classes' discrimination are mainly due to the differences in the spatial resolutions of the images acquired by both sensors (500 m for MODIS and 300 m for MERIS). The land cover classes that frequently mix at MODIS spatial resolution are those that have contiguous spatial distributions at the national scale and frequently have overlapping borders in the study area (e.g. grasslands and shrublands; mix forest stands and pure forest stands). Therefore, those land cover classes are generally better depicted by MERIS data because the geographic areas they occupy are generally smaller than the spatial resolution of MODIS images. Moreover, we verified that land cover classes that are constituted by the same cover type elements, but in different area proportions (e.g. open forest stands and grasslands), are as well better discriminated with MERIS images. We believe that in this situation the spectral information of MERIS images is prevailing for the discrimination of these vegetated land cover classes, namely due to the spectral wavelength data related with the "vegetation's red edge".

Now, looking at the classification strategy, we used a methodological approach that combined the classification of static land cover types with a Linear Discriminant Classifier (LDC) and the classification of intra-annual dynamic land cover types with a vegetation index differentiating technique. In Chapter 5, we found that by reducing the remotely-sensed data set through LDC, it is possible to achieve overall and specific land cover classification accuracies that are quite similar to those attained with recently proposed machine learning approaches, such as the Support Vector Machines (SVM). This holds true as long as the land cover nomenclature and the training sample are properly delineated for the landscape characteristics of the study area and used satellite images.

In this study, we showed that the SVM does not clearly overcome more standard parametric and non-parametric classifiers. The machine learning only performed better than the other competing algorithms when using high dimensional input multispectral remotely-sensed data sets and training sample sizes of reduced dimension. Alternatively, LDC does not need to be tuned and the user interaction in the process is minimal, thus reducing the inherent subjectivity associated with the SVM classifier. Indeed, we found that SVM are extremely difficult to tune for the optimal definition of the hyper-planes of separation between the land cover classes of interest.

In Chapter 6, it was also demonstrated that the Self-Organizing Maps (SOM) are not as good as LDC for the land cover classification problem that we wish to resolve. On the one hand, the classification accuracy attained with the SOM is lower than that attained with LDC, and on the other hand, the complexity inherent to the training of the SOM is not practicable on a regular mapping production context. Therefore, from an operational viewpoint, the high overall and per class thematic accuracies proved that LDC is a good option for the automatic land cover classification task, as long as the other components of the mapping protocol are correctly defined.

We now would like to elaborate on the sampling design that we applied for the computation of accuracy point estimators, absolute precisions estimates and for the testing sample collection. It is important to insist on the fact that the validation of land cover products must be based on a random sample collected through the use of some probabilistic sound design. This is imperative to correctly publish the research results to the science community. It is easy to carry out an accuracy assessment process like this and it should be implemented in each land cover mapping program. In this study, we proposed the use of a stratified sampling design to estimate the overall and per class accuracies of our land cover map (Chapters 7 and 9) and to collect the testing sample (Chapter 8). Stratified sampling must be used to collect

the testing sample when there is a need to ensure a minimum sample size in each stratum, i.e. to derive specific absolute precision estimates for all land cover classes presented in the map. In our case, the goal was to achieve absolute precisions inferior to 10% for all classes (Chapters 7, 8 and 9). Because the strata sizes were known, i.e. the areas occupied per land cover class in the map, the stratified sampling design estimators were used also to derive an overall accuracy absolute precision estimate more precise than would be obtained from simple random or systematic sampling without stratification. We achieved an absolute precision of only 2% for map overall accuracy (Chapter 9). Moreover, there was also a gain in the estimated overall accuracy because the contribution of accuracy estimates from the individual classes was weighted according to the area occupied per each class in the map.

Our assessment strategy improved the precision of map overall accuracy estimate and guaranteed that the precision of accuracy estimates were adequate for all land cover classes, even for those occupying a small proportion in the Portuguese territory. It is important to mention that the time and human labor used to implement this accuracy assessment protocol were smaller than those that would have been necessary to attain the same results by using a different sampling design.

11.2. Main difficulties

The outcomes of the different stand-alone manuscripts, presented in this dissertation, were derived from experimental research works performed with satellite images acquired over the Portuguese territory. The land cover classes of the nomenclature are well adapted to the landscape characteristics of Continental Portugal at the national scale. Yet, the proposed methodological approaches are valid and can be used for other nomenclatures, territories and image data sets. However, it is important to mention that the biophysical conditions that characterize the Mediterranean landscape of the Portuguese territory influence the time contribution of spectral information for land cover classes' discrimination. Therefore, the ordering of the spectral bands and dates that was established for the automatic land cover classification in this region can not be used straightforwardly within other regions.

It is also important to remark that there are differences between the samples used for testing the classifiers performances, and those used to assess the thematic accuracy of land cover maps. To train the classifiers and compare their individual performances, we resort to a sample of image pixels per land cover class that was deterministically selected. To evaluate the thematic accuracy of land cover maps, we selected a random sample collected all over the mainland territory using a stratified sampling design. The standard differences in the samples collection were established along with the goals of each specific study. Therefore,

because we used different samples within different research works, it is expected that some of the final accuracy results do not match between the stand-alone manuscripts.

Another limitation faced during this research is related to the quality of the previously quoted reference sample data sets. Note that these samples were used for the analysis of land cover classes discrimination, classifiers evaluation and selection, and for the accuracy assessment of the derived land cover maps. We can assert that the whole research is grounded on the sample data sets that were collected by visual interpretation of very high spatial resolution imagery, namely orthorectified aerial images with four spectral bands covering the whole Portuguese territory. The hypothetical occurrence of errors during the labeling of observations can bias the outcomes of the several studies. Indeed, there are at least two factors that may affect the quality of the training and testing sample databases: 1) the misregistration between the aerial images and the medium spatial resolution satellite images used in the different studies; and 2) the subjectivity inherent to the labeling of the reference sample observations by visual interpretation of the aerial images. Nevertheless, the medium spatial resolution satellite images were geometrically corrected to match as good as possible the orthorectified aerial images, and the reference sample database was recursively collected by several interpreters to reduce the subjectivity inherent to the particular visual interpretations and labeling of land cover classes' observations.

Because this research was focused mainly on the analysis of time series sets of multispectral satellites images, the probability of atmospheric noise incidence in the data sets is high. Moreover, observations of the same land cover class can be corrupted by atmospheric noise at different times and this also disturbs the regular classification process. In addition, because we focus on the classification of the whole Portuguese territory, it is obvious that individuals of the same class that are located at distant geographic locations will certainly be affected by different atmospheric disturbances. This is a severe problem that needed to be corrected in order to tune the intra-annual multispectral time profiles of land cover classes for posterior analyses. Therefore, to identify and remove the most important disturbances in the images data sets, we used the MODIS and MERIS surface reflectance quality assessment layers. Nevertheless, it is predictable, and it was verified in practice, that some minor radiometric errors are not identified by the quality assessment layers and still remain in the data sets. If this noise is not removed, then the land cover classification results attained with multitemporal datasets can be biased due to time-specific contaminations of classes' observations. As these disturbances are similar in kind to a shot noise superimposed to the relevant signal, we tried to correct them through the use of smoothing techniques (Chapters 2 and 3). However, these corrections on the observed time series may introduce some new

unknown errors in the real time series and thus bias the performed analyses (see Chapter 10). A good alternative to tackle this issue is the nonlinear harmonic model we presented in Chapter 10. However, because of the research and manuscripts publication timings, it was not possible to use the model for the time series correction tasks performed in Chapters 2 and 3.

We would like to remark that we always performed a global assessment of the contribution of the temporal and spectral satellites images information for land cover classes' discrimination. We did not investigate the differences in land cover classification accuracy at the individual level for different experimental conditions. Therefore, it is possible that some interesting and important outcomes were not totally revealed during the several experimental trials. Indeed, the McNemar test can be more insightful than the simple standard analysis of the difference between two proportions to compare the accuracy results of two classifiers that have been run on the same testing data.

11.3. Recommendations and future work

At the end of this dissertation, we would like to point out some recommendations about the automatic mapping of land cover distribution and the analysis of land cover time evolution using multispectral time series sets of medium spatial resolution satellite images. Moreover, we would like to indicate some future research work that may improve the outstanding results that were already attained during this investigation.

In the framework of an automatic classification protocol for mapping the land cover distribution from satellite images, we guess that the classifier is as much important as the other parts involved in the process. Because each classifier has its own specificities that restrict the framework of its application, there is no single "best" classifier. Indeed, the same classifier applied to different problems and trained with different data performs differently. There are other components in a practical work like this that also have an important part in the whole process. For example, it is extremely important to perform a deep analysis of the landscape characteristics and select the most adequate land cover nomenclature to represent them. It is impossible to distinguish between landscape features whose extent lie below the spatial resolution of the used satellite images. It is extremely difficult to distinguish between land cover types that are constituted by similar land cover features. Thus, land cover maps producers must be reasonable and do not expect to automatically distinguish between land cover features that are not separable within the spatial, spectral and temporal resolutions of the used satellite images. Indeed, they must recognize that many difficulties in the

classification process are associated to numerous factors that do not depend only on the used classifier.

Furthermore, the classification approach can also be more important than the classifier itself in the whole mapping process. The classifier does not need to be innovative or perfect (note that there is no classifier like this!); standard classifiers work well under clear classification approaches! Instead of looking for the best classifier, we should look in the first place for the best classification strategy. This is the case, e.g. of the classification strategy that we used to distinguish between intra-annual static and dynamic land cover classes (Chapter 9). We used different classifiers to map each of those classes and then find a simple way to combine the outputs of each classification into a unique final map. We suggest the producers to focus on these analyses and split the classification process into simple tasks, which results can be merged at a post-classification stage.

The reference land cover samples collected for training and testing the classifier are the basis for the whole work. However, producers discard the importance of this task and often the land cover sample observations are not properly collected because: the ancillary information used for the reference sample collection (e.g. high resolution aerial images) is not contemporaneous with the satellite images used for the automatic classification; the spatial resolution of ancillary information is not fine enough for visual identification of the correct land cover observation label; the image interpreters are not experts in discerning land cover types through visual analysis of remotely-sensed images. These difficulties may introduce noise in the sample, change the land cover classes' statistics and penalize the effective classification results, which are independent of the used classifier. Therefore, one suggests the producers to give special attention to these difficulties instead of trying to augment the classification accuracy by using the same training sample and simply changing the classifier.

We must insist on the fact that the collected testing samples are simple representations of the truth and not the real reference data. Accordingly, the disparities between the classification results and the sample observations can occur due to errors in the sample database and not due to the more usual classification errors. Therefore, to enhance the accuracy assessment of land cover maps, we suggest the use of confidence ratings collected during the labeling of sample observations. This approach was presented in Chapter 8 and the goal is to complement the traditional map evaluations. The confidence ratings attributed by the image interpreters to the observations labels serve as input for the development of assessment techniques based on fuzzy sets.

For future research, we recommend the use of mosaic classes in the nomenclature. These classes, such as agriculture with natural areas, represent independent landscape features that are characteristic of the Portuguese landscape at the spatial resolutions of MODIS and MERIS images. Indeed, these classes are included in the nomenclature that we proposed for mapping the study area (Chapter 8), but they were disregarded during the classification process. The problem is that these classes can not be classified as similar as the remaining intra-annual static land cover classes of the nomenclature. Because the pixels of mosaic classes contain small area features of several homogeneous land cover types, it is likely that a standard discrete classifier will automatically label the individuals of those classes in one of the homogeneous land cover types and vice-versa. We did not investigate this situation, but we think that the use of sub-pixel analyses, fuzzy techniques and especially of continuous classification schemes (e.g. regression tree algorithm) that categorize the percentage of different homogeneous land cover types within each pixel, are excellent approaches to deal with this issue.

During this study it was not possible to combine the optical spectral wavelength data, acquired by MERIS and MODIS, with other remotely-sensed data, such as those acquired by radar systems. In the past years, this task was not so easy because these data were acquired by different satellites, at different times, with different spatial resolutions, and with different standards of geometric correction. Therefore, the harmonization of both types of spectral wavelength data was difficult and clearly the classification results would not benefit from this union. Nowadays, the ENVISAT satellite carries a radar system onboard (i.e. ASAR), that acquires spectral wavelength data with the same spatial and temporal resolutions of MERIS images, and with the same geometric correction standards. Therefore, the use of radar data for land cover mapping and its fusion with these medium spatial resolution optical images is now easier and promotes the development of research studies on this subject.

We suggest that the nonlinear harmonic model can be effectively adjusted for the identification and prediction of environmental disturbances that harmfully modify the natural ecosystems. For example, in Portugal, we can recall a couple of challenging issues that may benefit from the use of the model-based prediction approach, e.g. the monitoring of the quantity and quality of habitats available for the reproduction of the Iberian lynx; and the monitoring of the extension of stubble fields from elder agricultural areas for the nesting of the great bustard.

To close this manuscript, we would like to summarize the key contributions of this investigation and to elaborate on the experience acquired during the PhD work. To do this

research we needed a problem, and to solve our problem we needed a motivation. From a general viewpoint, we wanted to turn possible and cost-effective the regular mapping of land cover distribution and the land cover monitoring at the national scale from satellite images. This is necessary to support numerous real-world applications, namely for forest fire prevention, natural resources assessment, crop yields estimation and carbon sink prediction. Current approaches to automatic land cover characterization from multispectral and multitemporal satellite images often fail due to the selected EO data, used data analysis methods and algorithms, and accuracy assessment protocols. In practice, we learned that the fast development of remote sensing technology goes beyond the methods and algorithms used for land cover characterization. Therefore, even resorting to recently proposed data analysis techniques, we are not able to cope with the massive EO data acquired by enhanced satellite sensors. To tackle this issue, our strategy was to simplify the data in order to serve as input features for current pattern recognition and machine learning methods. We proposed and tested two different approaches for two different applications: to perform an extraction of the most important information from satellite images for land cover mapping; to use a parsimonious model to compress into few informative features the whole information contained in the enhanced satellite images data sets for the identification and prediction of phenological attributes. Both approaches were applied in real case studies and revealed successful, thus promoting their future applications in operational programs.

This dissertation was not intended as a comprehensive survey of the state of the art of the design and analysis procedures for land cover products accuracy assessment. However, during this investigation, we confirmed that the evaluations of land cover products that are published in the literature are either insufficient or incorrectly reported. Despite the widespread acceptance of accuracy assessment, the basic structures of a statistically sound accuracy assessment are not recognized. Our goal was to discuss options available for land cover products evaluation before being used in scientific investigations and policy decisions. Moreover, we recommended the use of a clear, statistically sound and adequate accuracy assessment strategy to support the utility of land cover products for diverse national applications.

Numerous research directions were always open during the course of this investigation and some are still open. The most difficult part in this research was to select the appropriated path to follow during the whole investigation. The number of open doors was immense after solving each particular and difficult issue. We suppose that many points were neither studied nor solved during this study. Therefore, we realized that the end of this investigation is just the start for a future line of research.