

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE ENGENHARIA GEOGRÁFICA, GEOFÍSICA E ENERGIA



**Análise multitemporal de imagens do satélite Sentinel-2 como
suporte à elegibilidade das ajudas comunitárias agrícolas**

Susana Vicêncio Moura Feio

Mestrado em Engenharia Geográfica

Trabalho de Projeto orientado por:
Prof. Dr. João Catalão Fernandes e Prof. Dr. Ana Navarro Ferreira

2017

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE ENGENHARIA GEOGRÁFICA, GEOFÍSICA E ENERGIA



**Análise multitemporal de imagens do satélite Sentinel-2 como
suporte à elegibilidade das ajudas comunitárias agrícolas**

Susana Vicêncio Moura Feio

Mestrado em Engenharia Geográfica

Trabalho de Projeto orientado por:
Prof. Dr. João Catalão Fernandes e Prof. Dr. Ana Navarro Ferreira

2017

Resumo

O projeto apresentado foi desenvolvido em parceria com o Instituto de Financiamento da Agricultura e Pescas (IFAP, I.P.) e consistiu no desenvolvimento de uma metodologia para facilitar o controlo agrícola feito atualmente através de fotointerpretação, propondo uma forma mais eficiente de validar as declarações dos agricultores que requerem ajudas monetárias. A zona de estudo foi sugerida pelo IFAP por conter culturas de interesse e pela sua proximidade, abrangendo os distritos de Santarém, Setúbal, Évora, Leiria e Lisboa. A informação sobre as culturas constante no Sistema de Identificação Parcelar (SIP) utilizada no projeto é relativa à campanha de pedidos de ajuda agrícola de 2016. A análise da série temporal de imagens engloba 13 imagens, adquiridas pelo satélite Sentinel-2A, entre abril e setembro de 2016. Assim, através de uma série temporal de imagens é possível obter um conjunto de informações sobre as várias culturas existentes na região. Neste estudo foram estudados 4 tipos de culturas, sendo elas arroz, tomate, milho e pastagens. Calcularam-se 9 índices de vegetação, sendo o principal o índice de vegetação de diferença normalizada (NDVI). Com o objetivo principal de validar as declarações, a metodologia desenvolvida permite identificar as parcelas suspeitas e que necessitam de controlo por poderem corresponder a culturas declaradas erradamente. A validação foi realizada através da caracterização de culturas essencialmente com a série temporal de NDVI, embora tenham sido analisadas séries temporais calculadas com os outros índices de vegetação. Para além disso, utilizando as diversas bandas das imagens representativas de cada uma das fases fenológicas das culturas, foi efetuada a classificação das culturas de interesse recorrendo a três classificadores: Máxima Verosimilhança, Redes Neurais e Distância de Mahalanobis. Os resultados obtidos permitiram distinguir as pastagens das outras culturas, devido ao seu comportamento linear na resposta do NDVI ao longo da série temporal, mas também permitiram estabelecer padrões para o comportamento das restantes culturas. Foi ainda possível controlar a existência de plantação nas parcelas controladas, excluindo as parcelas que apresentavam um comportamento diferente do esperado e calcular a área ocupada por cada cultura com base no melhor resultado obtido da classificação das imagens.

Palavras-chave: Deteção Remota, NDVI, IFAP, Culturas, Parcelas.

Abstract

This project was developed in partnership with the Instituto de Financiamento da Agricultura e Pescas (IFAP, I.P.) and consisted in the development of a methodology to facilitate the agricultural subsidy control done currently by photo interpretation, proposing a more efficient methodology to validate the farmers' area-based financial support. The study area was suggested by IFAP because it contains cultures of interest, including Santarém, Setubal, Évora, Leiria and Lisbon districts. The information about crops used in this project was retrieved from the National Land Parcel Identification System (SIP) and is related to the subsidy control campaign of 2016. Time series analysis includes 13 images, acquired by the Sentinel-2A satellite, between April and September 2016. With the use of images time series, it is possible to obtain a set of information about the different cultures in the region. In this study 4 types of crops considered, rice, tomatoes, corn and pastures. Using the images, 9 vegetation indexes were calculated, being the Normalized Difference Vegetation Index (NDVI) the most relevant. With the main objective of validating the farmers' declarations, the developed methodology allows to identify the suspicious parcels that need to be field validated, since they don't correspond to the declared culture. The validation was performed through the characterization of crops, mainly with the NDVI time series, although other vegetation indices time series have been analyzed. In addition, several bands of the images that characterize each stage of the crop's growth cycle, were used as input for classification applying 3 different classifiers: Maximum Likelihood, Neural Network and Mahalanobis Distance. The results obtained enable to distinguish pastures from other cultures, due to its linear behavior in NDVI response along the crops' development cycle, but also enable to establish patterns for the other cultures behavior. Furthermore, it was still possible to control the existence of a crop in the controlled parcels, excluding the parcels that showed a different behavior than the one expected for that crop and to calculate the area occupied by each crop based on the best result obtained from the images' classification.

Keywords: Remote Detection, NDVI, IFAP, Cultures, Parcels.

Agradecimentos

Em primeiro lugar, aos meus pais por todo o incentivo e apoio incondicional demonstrado ao longo destes anos e por nunca me terem deixado desistir nos momentos mais difíceis e por terem sempre acreditado mesmo quando eu duvidei que era possível. Também à minha irmã por sempre ter acreditado em mim.

À Prof. Dra. Ana Navarro e ao Prof. Dr. João Catalão, orientadores deste projeto, toda a disponibilidade prestada para a resolução de problemas e revisão de conhecimentos durante a realização deste trabalho.

A todos os professores do curso de Engenharia Geográfica, obrigada por todo o apoio e pelos momentos de convívio aos quais sempre que possível estiveram presentes.

À Eng. Célia Sousa, ao João Correia e a toda equipa do parcelário do IFAP, o apoio e a cedência de informação necessária para a elaboração deste projeto.

Aos meus colegas e amigos, um grande obrigada por todas as horas e por todos os bons momentos passados durante esta fase, tornando este caminho mais fácil.

Em especial, Ana Catarina Ricardo, André Teles, António Morgado, Bruno Rita, Carolina Rocha, Catarina Rodrigues, Filipa Inácio, Francisco Santos, Gonçalo Nunes, Inês Martins, Joana Contente, João Brito, José Pedro Godinho, Luís Ribeiro, Mariana Melo, Maria João Martins, Tiago Antunes e Vasco Laranjinha.

Índice

1. Introdução	1
1.1 Enquadramento e Motivação	1
1.2 Objetivos.....	2
1.3 Estado da arte.....	2
1.4 Estrutura da tese.....	5
2. Metodologia e dados	7
2.1 Área de estudo	7
2.2 Dados	8
2.3 Metodologia.....	11
2.3.1 Cálculo dos índices de vegetação	12
2.3.2 Classificação de imagens	13
3. Resultados	17
3.1 Índices de vegetação	17
3.2 Classificação de imagens	23
3.2.1 Classificador de Máxima Verossimilhança	24
3.2.2 Classificador de Redes Neurais	26
3.2.3 Classificador de Distância de Mahalanobis	29
3.3 Validação de declarações.....	30
4. Conclusões	34
Referências.....	36
Anexos	38

Índice de Figuras

Figura 2.1: Área de estudo	08
Figura 2.2: Pormenor da área de estudo que compreende a lezíria do Ribatejo	07
Figura 2.3: Parcelas em estudo	10
Figura 2.4: Fluxograma da metodologia utilizada	11
Figura 2.5: Combinação das 4 bandas para as três imagens utilizadas no estudo	14
Figura 2.6: Combinação das 7 bandas para as três imagens utilizadas no estudo	14
Figura 2.7: Combinação das 9 bandas para as três imagens utilizadas no estudo	14
Figura 2.8: Combinação dos índices utilizados para o estudo	15
Figura 3.1: Imagens NDVI para zona de cultura de Milho Tardio	18
Figura 3.2: NDVI Milho (esquerda) e Tomate (direita)	18
Figura 3.3: NDVI Arroz (esquerda) e Pastagens (direita)	19
Figura 3.4: NDVI Tomate Precoce (esquerda) e Tomate Tardio (direita)	20
Figura 3.5: NDVI Milho Precoce (esquerda) e Milho Tardio (direita)	20
Figura 3.6: NDVI Arroz (esquerda) e Elemento Linear de Arroz (direita)	20
Figura 3.7: NDVI Pastagens Permanentes (esquerda) e Pastagens Arbustivas (direita)	21
Figura 3.8: CIREdEdge para a cultura de tomate (esquerda), milho (centro) e arroz (direita)	21
Figura 3.9: GNDVI para a cultura de tomate (esquerda), milho (centro) e arroz (direita)	22
Figura 3.10: Gráfico com a média para cada cultura ao longo da série temporal	23
Figura 3.11: Exemplos de pormenores de zonas com parcelas de milho e tomate	31
Figura 3.12: Exemplos de pormenores de zonas com parcelas de arroz e pastagens	32
Figura Anexo 1: Imagens com 4 bandas para as 4 datas escolhidas para a validação com 4 classes utilizando o Classificador de Máxima Verosimilhança	38
Figura Anexo 2: Imagens com 7 bandas para as 4 datas escolhidas para a validação com 4 classes utilizando o Classificador de Máxima Verosimilhança	39
Figura Anexo 3: Imagens com 9 bandas para as 4 datas escolhidas para a validação com 4 classes utilizando o Classificador de Máxima Verosimilhança	40
Figura Anexo 4: Imagens com 8 índices para as 4 datas escolhidas para a validação com 4 classes utilizando o Classificador de Máxima Verosimilhança	41

Figura Anexo 5: Imagens com as 4 bandas para as 4 datas escolhidas para a validação com 8 classes utilizando o Classificador de Redes Neurais.....	42
Figura Anexo 6: Imagens com as 4 bandas para as 4 datas escolhidas para a validação com 6 classes utilizando o Classificador de Redes Neurais.....	43
Figura Anexo 7: Imagens com as 4 bandas para as 4 datas escolhidas para a validação com 4 classes utilizando o Classificador de Redes Neurais	44
Figura Anexo 8: Imagens para o índice NDVI	45
Figura Anexo 9: Imagens para o índice CIREdEdge	46
Figura Anexo 10: Imagens para o índice CIGreen	47
Figura Anexo 11: Imagens para o índice NDI7	48
Figura Anexo 12: Imagens para o índice NDI5	49
Figura Anexo 13: Imagens para o índice NDTI	50
Figura Anexo 14: Imagens para o índice VIGreen	51
Figura Anexo 15: Imagens para o índice NDSVI	52
Figura Anexo 16: Imagens para o índice GNDVI	53
Figura Anexo 17: Zoom de zonas com maior incidência de cada cultura	54
Figura Anexo 18: Percentagem de cada cultura existente no estudo	54

Índice de Tabelas

Tabela 2.1: Ordem cronológica das datas das imagens de satélite utilizadas para o estudo.....	09
Tabela 2.2: Resolução espacial das bandas das imagens	09
Tabela 2.3: Percentagem ocupada pelas culturas	10
Tabela 2.4: Índices de vegetação	12
Tabela 3.1: Controlo das parcelas controladas fisicamente	22
Tabela 3.2: Resultados da exatidão global para as 4 datas e imagem de 4 bandas utilizando a Máxima Verosimilhança	24
Tabela 3.3: Resultados estatísticos para as imagens das 9 bandas e dos índices de 19 julho 2016	25
Tabela 3.4: Precisão Global para as 4 datas escolhidas para a validação	25
Tabela 3.5: Resultados estatísticos para as imagens das 9 bandas e dos índices de 19 julho 2016	26
Tabela 3.6: Comparação de resultados estatísticos com 8 classes para as imagens das bandas e dos índices de 19 julho 2016	27
Tabela 3.7: Comparação de resultados estatísticos com 6 classes para as imagens das bandas e dos índices de 19 julho 2016	28
Tabela 3.8: Comparação de resultados estatísticos com 8 classes para as imagens das bandas e dos índices de 19 julho 2016	29
Tabela 3.9: Comparação de resultados estatísticos com 8 classes para as imagens das bandas e dos índices de 19 julho 2016	30
Tabela 3.10: Percentagem de parcelas válidas e parcelas suspeitas para cada cultura	31

Índice de Anexos

Anexo I – Resultados do Classificador de Máxima Verosimilhança	38
Anexo II – Resultados do Classificador de Redes Neurais	42
Anexo III – Índices de vegetação	45
Anexo IV – Parcelas das culturas	54

Lista de Abreviaturas

IFAP – Instituto de Financiamento da Agricultura e Pescas, I.P.

DR – Detecção Remota

DA – Detecção de alterações

NDVI – Normalized Difference Vegetation Index

NDSVI – Normalized Differential Senescent VI

GNDVI – Green Normalized Difference Vegetation Index

NDI7 – Normalized Difference Index 7

NDI5 – Normalized Difference Index 5

ViGreen – Green Vegetation Index

CiRedEdge – Red-Edge Chlorophyll Index

CiGreen – Green Chlorophyll Index

SWIR – Short-wavelength infrared

1. Introdução

1.1 Enquadramento e Motivação

O presente projeto foi desenvolvido em colaboração com o Instituto de Financiamento da Agricultura e Pescas (IFAP, I.P.) no âmbito das suas atividades de controlo às ajudas comunitárias ao sector agrícola.

O IFAP, I.P., tem por missão proceder à validação e ao pagamento decorrente do financiamento da aplicação de diversas medidas definidas a nível nacional e comunitário, no âmbito da agricultura, desenvolvimento rural, pescas e setores conexos, bem como propor as políticas e estratégias de tecnologias de informação e comunicação no âmbito da agricultura e pescas. O IFAP tem como atribuições:

- Garantir o funcionamento dos sistemas de apoio e ajudas diretas nacionais e comunitárias e a aplicação, a nível nacional, das regras comuns para os regimes de apoio direto no âmbito da política comum;
- Garantir o cumprimento da função de organismo pagador do Fundo Europeu Agrícola de Garantia (FEAGA) e do Fundo Europeu Agrícola de Desenvolvimento Rural (FEADER);
- Garantir o cumprimento da função de autoridade de certificação no âmbito do Fundo Europeu das Pescas (FEP), bem como de organismo intermédio na aceção do Reg. (CE) nº 1198/2006, de 27 de julho;
- Executar a política estratégica na área das tecnologias de informação e comunicação, para o setor da agricultura e pescas, assegurando a construção, gestão e operação das infraestruturas na respetiva área de atuação;
- Apoiar o desenvolvimento da agricultura e das pescas, bem como do setor agroalimentar, através de sistemas de financiamento direto e indireto.

Com a adesão da União Europeia (UE) ao sistema de subsídios baseados em áreas, foi desenvolvido pelo IFAP, o Sistema de Identificação Parcelar (SIP). O SIP é um sistema de informação geográfica que tem como objetivo a identificação do limite das parcelas das explorações agrícolas, às quais é atribuído um número único, assim como a delimitação e classificação das ocupações de solo, permitindo a apresentação de candidaturas a Ajudas Comunitárias e a execução de Ações de Controlo [IFAP, I.P., 2017].

No âmbito das suas atribuições, o IFAP realiza anualmente o controlo das ajudas aos agricultores mediante visitas às parcelas agrícolas com pedidos de suporte financeiro. Estas visitas realizam-se com o objetivo de verificar o cumprimento dos requisitos de concessão das ajudas e das normas aplicáveis no âmbito da condicionalidade. Neste controlo físico confirmam-se os limites das parcelas declaradas assim como as ocupações culturais encontradas e respetiva elegibilidade. São registadas as anomalias consideradas pertinentes e necessárias para a ação de controlo, ficando esta concluída depois de todos os intervenientes na ação de campo assinarem o relatório de controlo e realizado o tratamento informático da mesma.

As deslocações efetuadas para o controlo físico das declarações dos agricultores requerem a mobilização de um elevado numero de técnicos, de equipamentos e veículos o que torna esta operação bastante onerosa para o IFAP. Nos últimos anos a Comunidade Europeia tem incentivado os países aderentes a implementarem procedimentos alternativos que reduzam os custos das operações de controlo físico. O controlo por Detecção Remota surge como uma das técnicas mais promissoras para redução de custos das operações de controlo,

sendo realizada pelo IFAP há mais de 20 anos. Contudo, o uso de imagens de satélite, pelo IFAP, é limitado à interpretação visual por foto intérpretes numa perspetiva de complementaridade ao controlo realizado no terreno. As imagens disponibilizadas pela Comunidade Europeia para este efeito são geralmente imagens de alta resolução dos satélites SPOT e GeoEye.

1.2 Objetivos

Neste projeto pretendemos desenvolver um sistema de apoio ao controlo das ajudas comunitárias à produção agrícola com recurso a imagens multiespectrais do satélite Sentinel-2. A abordagem consiste na análise de séries temporais de imagens multiespectrais que permita caracterizar o ciclo fenológico das culturas. As imagens multiespectrais possibilitam o cálculo de vários índices vegetativos combinando as bandas do visível (VIS) e do infravermelho próximo (NIR) que permitem caracterizar o ciclo vegetativo da cultura. A utilização das bandas SWIR (*Short-wavelength infrared*) e das bandas no limiar do vermelho (*Red-Edge*) também foi considerada para o cálculo de outros índices, para avaliar o seu potencial para uma melhor discriminação entre diferentes culturas agrícolas.

1.3 Estado da arte

A natureza dinâmica da agricultura, como a sazonalidade e a sua ocorrência em quase todo o globo, fazem parte do forte incentivo para os cientistas monitorizarem a agricultura a partir do espaço. O controlo agrícola, com recurso à DR, é feito através de fotointerpretação. A DR é um meio que oferece informação, através da resposta espectral, sobre o estado da vegetação e a sua fenologia, medida e monitorizada através de sensores multiespectrais [CCRS, 2017]. Com a DR é possível identificar o tipo de cultura e a área de cada parcela evitando que seja necessário a deslocação ao local para controlo físico e obtenção, método utilizado para validação de declarações dos agricultores no IFAP.

Os pedidos de ajuda dos agricultores têm de ser obrigatoriamente controlados para não existirem benefícios de subsídios quando as parcelas não são exploradas segundo o declarado. O controlo destes pedidos pode ser feito de forma clássica, através de deslocações ao terreno, ou com a aplicação da DR, sendo esta uma ferramenta eficaz para fornecer informação de alta qualidade sobre a agricultura, providenciando dados objetivos e homogêneos que podem ser geograficamente e temporalmente registados [Tsiligrades, 1998].

As características espectrais do coberto vegetal existente na superfície terrestre variam consideravelmente no tempo e no espaço. A refletância espectral da vegetação apresenta uma grande variação em função das estações do ano, da espécie vegetal, do teor de água na planta e do comprimento de onda do sensor [Fonseca e Fernandes, 2004]. A clorofila, composto químico presente nas folhas, é um fator condicionante da resposta espectral, pois esta absorve fortemente os comprimentos de onda do vermelho e azul. Quando o teor de clorofila da vegetação está no máximo existe uma maior absorção dos comprimentos de onda vermelhos e maior reflexão de infravermelho. Por outro lado, quando há menos clorofila nas folhas, existe uma menor absorção e proporcionalmente uma maior reflexão dos comprimentos de onda vermelhos [CCRS, 2017].

A interpretação visual das imagens é o primeiro passo na análise de imagens de satélite. O reconhecimento de alvos é a chave para a interpretação e obtenção de informação. Visualmente podem detetar-se diferenças através do padrão, textura, forma ou tom de cor. Sendo as imagens analisadas visualmente banda a banda

ou combinando bandas, os diferentes tipos de ocupação do solo muitas vezes diferem entre si por comparação entre bandas espectrais [Sousa e Silva, 2011].

Diversos autores têm constatado que a utilização de imagens adquiridas em várias datas possibilita uma melhoria significativa na classificação de culturas, remetendo para a importância de análises de séries temporais neste tipo de estudos [Karjalainen *et al.*, 2008; McNairn *et al.*, 2009, Long *et al.*, 2013]. Long *et al.* [2013], referem que os dados multitemporais produzem melhores resultados que dados de uma única data, devido a estes registarem a diversidade espectral observada ao longo do ciclo fenológico da planta.

A classificação de culturas é então um fator chave na monitorização agrícola, que beneficia da análise de séries temporais de imagens correspondentes a um período de crescimento, utilizando o comportamento fenológico distinto de diferentes culturas [Zilmann e Weichelt, 2014]. Karjalainen *et al.* [2008] verificaram uma melhoria na exatidão global através da adição cumulativa de novas imagens de datas distintas ao processo de classificação, tendo obtido resultados com uma exatidão superior a 84%.

Serra e Pons [2008] apresentaram uma metodologia para mapear e monitorizar assinaturas espectrais de seis principais culturas mediterrâneas, cereais, arroz, milho, alfafa e frutas de árvores, principalmente, usando 36 imagens Landsat obtidas entre 2002 e 2005, tendo concluído que uma abordagem multitemporal é essencial para obter bons resultados. Simonneaux *et al.* [2008] também usaram uma série temporal de 8 imagens Landsat para identificar 4 tipos de classes, solo nu, vegetação, árvores em solo nu e árvores com vegetação à volta, usando um algoritmo com recurso a árvores de decisão.

Os índices de vegetação baseiam-se nas diferenças de refletividade que a vegetação de cor verde apresenta nas regiões do visível e do infravermelho. Têm dois objetivos principais, realçar aspetos que permitem discriminar entre diferentes tipos e estados de vegetação e minimizar radiâncias do mesmo tipo de coberto, resultantes de diferenças de iluminação devidas à topografia [Fonseca e Fernandes, 2004]. Os índices de vegetação fornecem informação sobre as condições da vegetação através da relação entre as bandas do satélite e as características das culturas. A comparação de 7 índices de vegetação aplicados a imagens de diferentes datas para detetar alterações na cobertura do solo levou a que Lyon *et al.* [1998] concluíssem que o índice de vegetação de diferença normalizada (NDVI) é o melhor para a deteção de alterações (DA) em zonas com vegetação.

O NDVI é muitas vezes utilizado para a monitorização global da vegetação, permitindo detetar as mudanças do coberto vegetal ao longo do tempo, em virtude de as razões de bandas compensarem as mudanças das condições de iluminação, do declive do terreno e exposição. As áreas de vegetação vão apresentar valores elevados devido às refletâncias relativamente elevadas no NIR (infravermelho próximo) e baixa refletância no *Red* (vermelho). Com valores negativos surgem as nuvens e a água, cuja refletância é maior no *Red* do que no NIR. Os solos têm refletâncias similares nas duas bandas e resultam em valores próximos de zero para os índices de vegetação [Sousa e Silva, 2011].

Outros índices de vegetação espectral têm sido amplamente usados para analisar e monitorizar variações temporais e espaciais nas características das culturas. Por exemplo, o NDVI e o índice de vegetação verde (VIGreen) foram usados para monitorizar as condições de vegetação e a estrutura da planta, mostrando diferentes respostas entre culturas com copa fechada ou copa aberta [Gitelson *et al.*, 2002; Jiang *et al.*, 2006]. Além disso, outros índices de vegetação que incorporam as bandas SWIR (infravermelho de ondas curtas),

tais como o NDSVI, o NDI7 e o NDTI, [Qi *et al.*, 2002; McNairn and Protz, 1993; van Deventer *et al.*, 1997], são importantes para a identificação do desempenho das culturas quando estas são avaliadas em relação às suas propriedades relacionadas com o estado do solo (solo nu), com a humidade, o vigor da vegetação e a fotossíntese presente.

A DA é definida por Singh [1989] como o processo de identificação de diferenças do estado de um objeto ou fenómeno através da sua observação em diferentes momentos, útil em diversas aplicações, como a análise da ocupação do solo, a monitorização agrícola, o estudo fenológico da vegetação, as mudanças sazonais, a deteção de stress na vegetação, a monitorização do degelo e outras mudanças ambientais.

O conhecimento das alterações na superfície torna-se, com o passar do tempo, cada vez mais importante na monitorização local, regional e global de recursos e do ambiente. A escolha de um conjunto de imagens do mesmo local, em datas diferentes, permite realizar um estudo de alterações, sendo que o método utilizado de entre todos os diversos métodos existentes, depende sempre da área em estudo e dos objetos do mesmo

A metodologia mais adotada em estudos de DA, caracterização e monitorização é a classificação baseada na comparação, tendo a grande vantagem de produzir uma matriz com a informação da alteração e a redução de impactos externos, tais como diferenças atmosféricas entre imagens multitemporais [Hussain *et al.*, 2013; Lu *et al.*, 2004]. Para esta metodologia é necessário um número suficiente de áreas de treino, com qualidade, de modo a que os resultados obtidos sejam o mais satisfatórios possível, sendo esta uma tarefa muitas vezes de grande dificuldade.

Para a classificação de culturas, as técnicas de DA podem basear-se em objetos, sendo obtidos melhores resultados comparativamente aos obtidos com classificações ao nível do píxel. Existem diversas técnicas de DA, sendo a mais usual a aplicação de um valor limite (*threshold*) para distinguir entre mudança e não mudança. O estabelecimento de um limiar constitui uma limitação, uma vez que um limiar demasiado baixo poderá excluir áreas que sofreram alterações, enquanto que um valor alto poderá incluir demasiadas áreas que sofreram alterações [Hussain *et al.*, 2013].

Os métodos mais simples consideram operações baseadas no píxel, enquanto que os mais complexos se baseiam em objetos. O primeiro passo da classificação de imagens consiste na elaboração de uma legenda apropriada com todas as classes de ocupação do solo necessárias ao estudo, contendo todas as possibilidades de tipos de ocupação [CCRS, 2017]. Existem vários métodos de classificação, que através de diversas abordagens procuram identificar com eficácia a informação de cada píxel ou objeto da imagem, classificando-o em categorias de acordo com a sua informação espectral.

O classificador de Máxima Verosimilhança é o mais utilizado em DR. É um classificador considerado paramétrico, pois envolve parâmetros, vetor da média e matriz de covariância, da distribuição gaussiana multivariada e é supervisionado, pois estima estes parâmetros através das amostras previamente definidas [Erbert, 2001]. Por sua vez, as Redes Neurais são um algoritmo mais moroso, com uma técnica não supervisionada avançada de redes neurais, constituída por unidades de processamento que aplicam funções matemáticas, logarítmica ou hiperbólica, capazes de reconhecer padrões e realizar atividades variadas aos dados recebidos [Pádua, 2000]. O classificador de Distância de Mahalanobis é um classificador de distância, sensível à direção, que usa estatísticas para cada classe, assumindo que todas as covariâncias da classe são iguais, tornando-se num método mais rápido que o de Máxima Verosimilhança [Harris, 2017].

O lançamento dos 2 satélites Sentinel-2, permite uma resolução espectral e espacial acrescida, com 13 bandas espectrais, que vão do visível e do infravermelho próximo, até ao infravermelho de ondas curtas, incluindo três bandas no limiar do vermelho, com diferentes resoluções espaciais que podem ir até aos 10m. Para além disso, as imagens destes satélites apresentam um tempo de revisita curto, de 5 dias, o que irá permitir a monitorização de fenómenos de evolução rápida, tal como o desenvolvimento de culturas agrícolas. Tais características proporcionam informação essencial sobre o estado da vegetação.

1.4 Estrutura da tese

Este projeto encontra-se dividido em quatro capítulos principais, sendo eles: (1) Introdução, (2) Metodologia e dados, (3) Resultados e (4) Conclusão. O primeiro capítulo faz referência à motivação e ao enquadramento deste projeto, mencionando quais os objetivos pretendidos para atingir os resultados propostos, descrevendo o estado da arte e a estrutura do documento. O segundo capítulo descreve os dados necessários para o estudo e a metodologia desenvolvida para atingir os objetivos, assim como a área de estudo. No terceiro capítulo são apresentados os resultados obtidos, sendo efetuada a análise e caracterização das culturas e por fim uma tentativa de validação das declarações dos agricultores. O quarto capítulo contempla a conclusão e sugestões para um possível trabalho futuro. Também existem alguns anexos, necessários para explicações mais detalhadas sobre o trabalho realizado.

2. Metodologia e dados

2.1 Área de estudo

A área de estudo foi proposta pelo IFAP por ser a área abrangida pelo controlo físico em 2016. Os distritos abrangidos foram Santarém, Setúbal, Évora, Leiria e Lisboa. A zona de estudo encontra-se delimitada pelos paralelos $39^{\circ}45'0.96''$ N e $38^{\circ}45'14.53''$ N e pelos meridianos $9^{\circ}0'0.84''$ W e $7^{\circ}44'11.75''$ W, como é possível ver na Figura 2.1, onde a exploração agrícola é intensiva. Sendo uma área predominantemente rural, é uma zona essencialmente plana exceto na zona nordeste. No período compreendido entre abril e setembro, é nos meses de verão que as temperaturas são mais elevadas, atingindo o seu pico no mês de agosto, onde rondam os 40°C , sendo a precipitação escassa nesta altura do ano, levando a um aumento de seca dos terrenos nesta zona, visível nos resultados obtidos. Na Figura 2.2 encontra-se um pormenor da área de estudo.

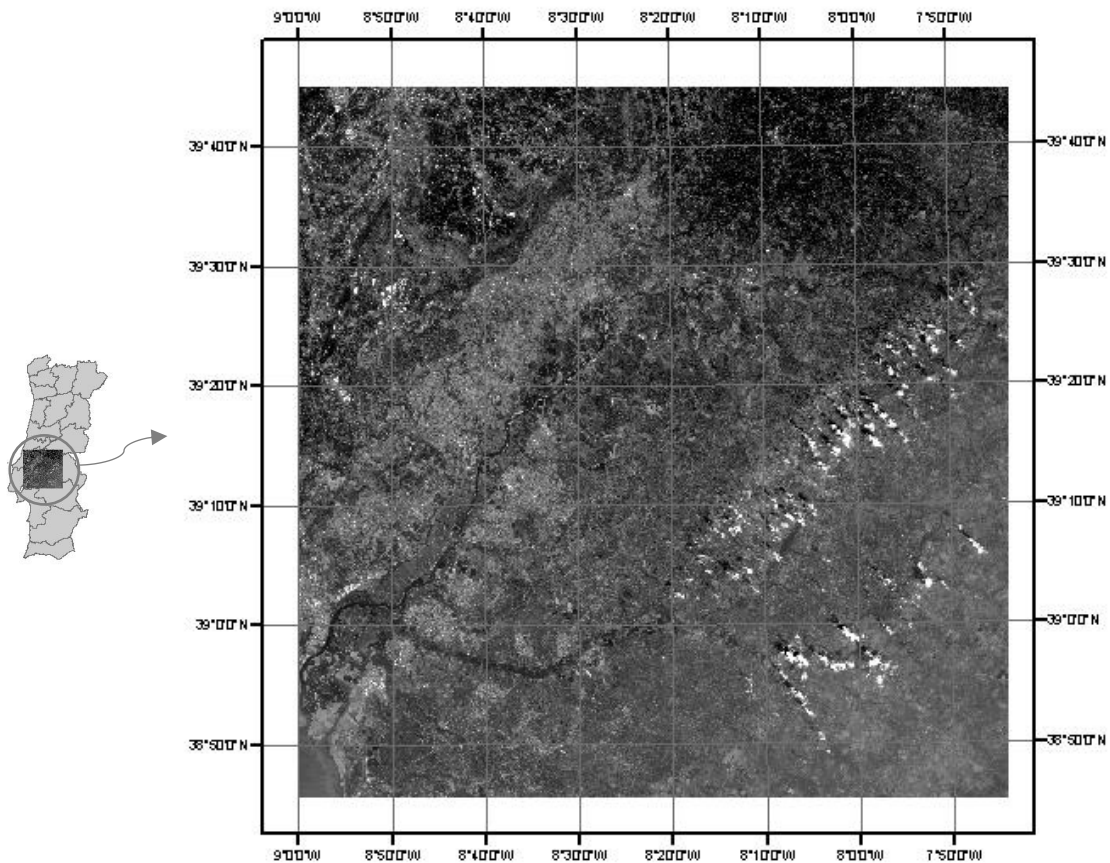


Figura 2.1: Área de estudo.

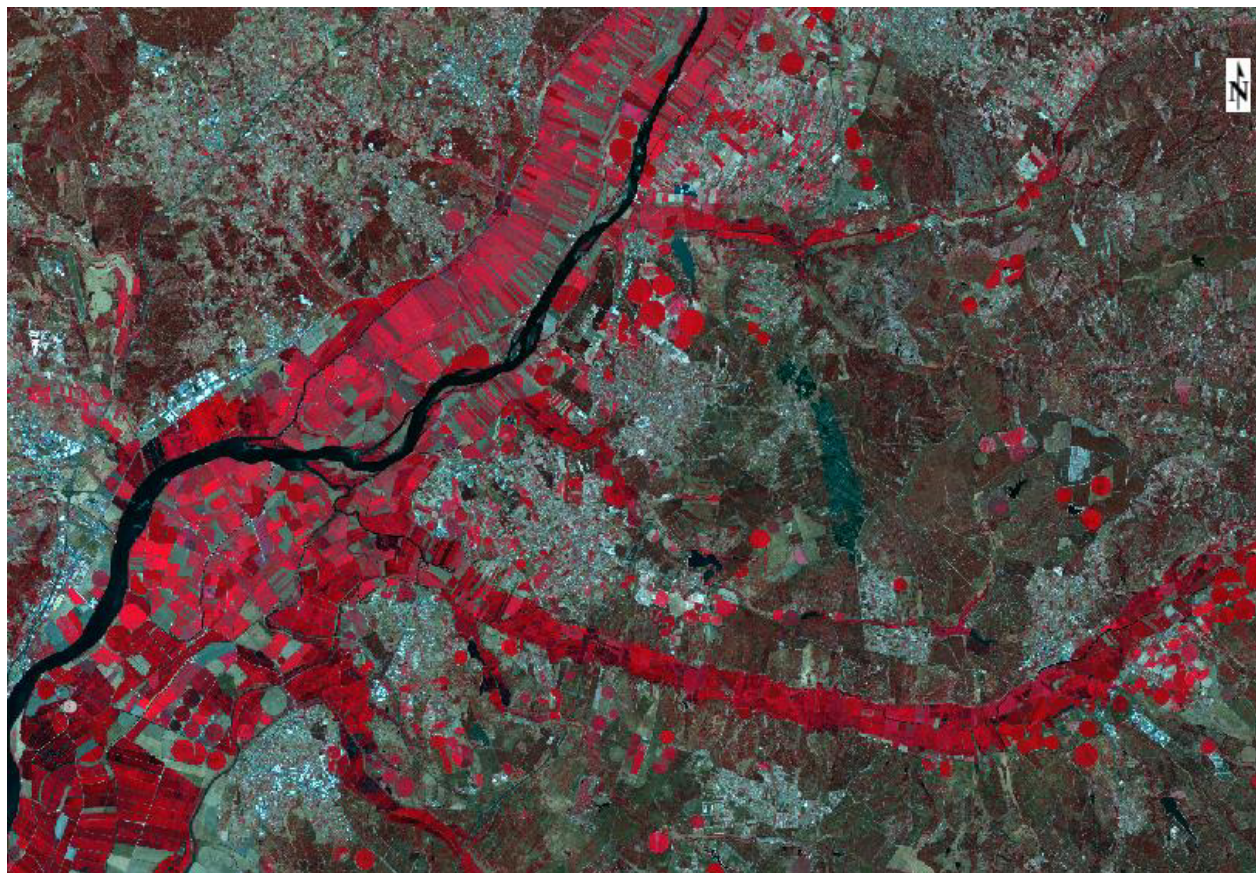


Figura 2.2: Pormenor da área de estudo que compreende a lezíria do Ribatejo.

2.2 Dados

As imagens utilizadas neste estudo foram adquiridas pelo satélite Sentinel-2. A missão Sentinel-2A tem uma órbita quase-polar e dispõe de um sensor MSI (*MultiSpectral Instrument*) com 13 bandas espectrais, com uma grande resolução espacial, de 10, 20 ou 60 metros, dependendo da banda, e uma resolução temporal de 10 dias, com um satélite, e de 5 dias com dois satélites operacionais. Cada imagem do satélite Sentinel-2 tem uma dimensão de 100 por 100 km, com 12 bits por píxel e uma resolução espacial de 10m. A missão Sentinel-2 garante a continuidade de aquisição de dados das missões SPOT e Landsat, destacando-se pela maior resolução espectral e espacial que apresenta [ESA, 2017].

Após consulta do site da ESA e analisadas as imagens adquiridas pelo Sentinel-2A no ano de 2016, foram selecionadas 13 imagens que reúnem as condições necessárias à realização do estudo multitemporal. As imagens não apresentam nuvens, não evidenciam efeitos atmosféricos (neblina e nevoeiro) e têm a mesma cobertura geográfica, ou seja, são sobreponíveis espacialmente. Na Tabela 2.1 estão listadas as datas das imagens.

Tabela 2.1: Ordem cronológica das datas das imagens de satélite utilizadas para o estudo.

Datas das imagens	
30 abril 2016	8 agosto 2016
20 maio 2016	18 agosto 2016
19 junho 2016	28 agosto 2016
29 junho 2016	7 setembro 2016
9 julho 2016	17 setembro 2016
19 julho 2016	27 setembro 2016
29 julho 2016	

Para cada uma das datas de aquisição selecionadas foram descarregadas as bandas B2, B3, B4, B5, B6, B7, B8, B11 e B12 (Tabela 2.2). As bandas foram posteriormente agrupadas numa estrutura BSQ (*Band sequential*), originando imagens com as combinações de bandas necessárias para o cálculo dos índices. Antes de agrupar, foi necessário efetuar a reamostragem das bandas de 20 m (B5, B6, B7, B11 e B12) para 10 m, para que assim toda a imagem ficasse com a mesma resolução geométrica.

Tabela 2.2: Resolução espacial das bandas das imagens [Fonte: ESA, [2017]].

Resolução	Nº Banda	Nome da Banda	Comprimento de onda (nm)
10 m	B2	<i>Blue</i>	490
	B3	<i>Green</i>	560
	B4	<i>Red</i>	665
	B8	NIR	842
20 m	B5	<i>Red-Edge 1</i>	705
	B6	<i>Red-Edge 2</i>	740
	B7	<i>Red-Edge 3</i>	783
	B11	Swir 1	1610
	B12	Swir 2	2190

Juntamente com as imagens, foi também utilizado um ficheiro vetorial do SIP com as delimitações das parcelas fornecido pelo IFAP, a partir do qual foi possível separar as diferentes culturas analisadas no estudo (Figura 2.3).



Figura 2.3: Parcelas em estudo (amarelo: milho, vermelho: tomate, roxo: arroz, verde: pastagens).

As parcelas, porção contínua de terreno homogêneo com limites estáveis agrônomicamente e geograficamente com uma identificação única, foram fornecidas num ficheiro vetorial em Datum 73 transformado para UTM/WGS84, sendo este o sistema em que se encontram as imagens de satélite. O ficheiro contém também um atributo em que é possível ver qual o tipo de cultura controlada fisicamente para cada parcela. A informação relativa à cultura controlada é usada para se criar um subconjunto com as culturas a serem analisadas: tomate, arroz, milho, pastagens (Tabela 2.3). A cultura com maior ocupação cultural é pastagens, sendo que a área ocupada por estas, em comparação com as restantes culturas, é superior. A sua distribuição também é mais evidenciada, enquanto que as culturas de arroz, milho e tomate se encontram mais concentradas na zona oeste. A segunda cultura com maior ocupação é o milho, seguindo-se do tomate e por último a cultura de arroz.

Tabela 2.3: Percentagem ocupada pelas culturas.

	Relação à imagem	Relação às culturas
Tomate	0.08%	19.19%
Arroz	0.06%	13.52%
Milho	0.12%	28.63%
Pastagens	0.17%	38.67%

2.3 Metodologia

Neste projeto pretende-se desenvolver uma metodologia que permita efetuar o controlo remoto das declarações dos agricultores que requerem subsídios, de forma a tornar mais eficiente a tarefa desenvolvida pelo IFAP para a validação das parcelas declaradas. Sendo o controlo efetuado presentemente com deslocações físicas ao local, este torna-se dispendioso e moroso. Com um controlo feito através de imagens de satélite, a tarefa torna-se bastante mais simples e rápida, envolvendo menos custos para o instituto.

A metodologia proposta consiste na análise da série temporal de vários índices de vegetação para, numa primeira fase, identificar de forma inequívoca se uma determinada parcela foi cultivada, e numa segunda fase, identificar o tipo de cultura. Para a identificação do tipo de cultura foram exploradas duas abordagens. A primeira consistiu na avaliação da possibilidade de caracterizar as culturas agrícolas com base no comportamento de diversos índices de vegetação ao longo do ciclo fenológico. A segunda, consistiu na avaliação da possibilidade de discriminar as várias culturas declaradas com recurso a algoritmos de classificação de imagem paramétricos (Máxima Verosimilhança e Distância de Mahalanobis) e não paramétricos (Redes Neurais).

No fluxograma da Figura 2.3 é apresentada a metodologia utilizada para a concretização dos objetivos propostos para este projeto.

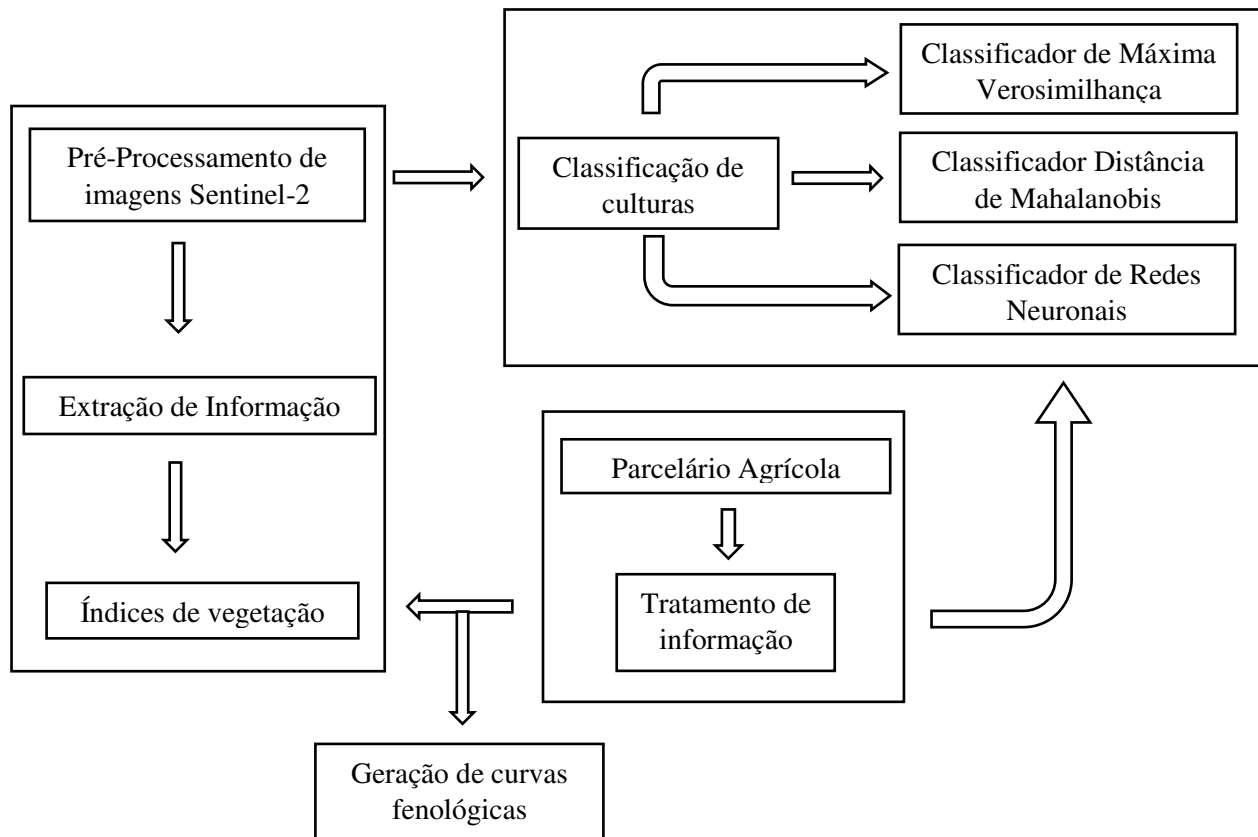


Figura 2.4: Fluxograma da metodologia utilizada.

Inicialmente começou-se por descarregar as imagens do Sentinel-2A das datas necessárias que abrangiam a zona coberta pelas parcelas cedidas pelo IFAP. Em seguida fez-se o tratamento das parcelas, analisando quais as parcelas correspondentes às culturas necessárias ao estudo e a sua distribuição geográfica. Posteriormente fez-se a análise gráfica através da resposta espectral das parcelas de cada cultura a cada índice na série temporal e também se procedeu à análise dos índices, combinando as bandas necessárias para cada um, criando as imagens respetivas e também as imagens com as combinações de bandas, assim como a classificação de imagens com o estudo de três conjuntos de classes e de três métodos de classificação, comparando os resultados.

2.3.1 Cálculo dos índices de vegetação

Foram considerados 9 índices de vegetação utilizados em diversos estudos para avaliar o seu potencial para melhorar a separabilidade espectral entre os diferentes tipos de culturas em estudo (Tabela 2.4).

Tabela 2.4 Diversos índices de vegetação utilizados neste estudo.

NDVI	$NDVI = \frac{B8 - B4}{B8 + B4}$	Rouse <i>et al.</i> , 1973
CIRedEdge	$CIRedEdge = \frac{B7}{B5} - 1$	Gitelson <i>et al.</i> , 2003
CIGreen	$CIGreen = \frac{B7}{B3} - 1$	Gitelson <i>et al.</i> , 2003
NDI7	$NDI7 = \frac{B8 - B12}{B8 + B12}$	McNairn e Protz, 1993
NDI5	$NDI5 = \frac{B8 - B11}{B8 + B11}$	McNairn e Protz, 1993
NDTI	$NDTI = \frac{B11 - B12}{B11 + B12}$	Van Deventer <i>et al.</i> , 1997
VIGreen	$VIGreen = \frac{B3 - B4}{B3 + B4}$	Gitelson <i>et al.</i> , 2002
NDSVI	$NDSVI = \frac{B11 - B4}{B11 + B4}$	Qi <i>et al.</i> , 2002
GNDVI	$GNDVI = \frac{B8 - B3}{B8 + B3}$	Gitelson e Merzlyak, 1996

Para cada cultura foi calculado o valor da média e do desvio padrão para cada um dos índices ao longo da série temporal, o que permitiu gerar gráficos com o ciclo de crescimento para cada cultura. Após o cálculo destes valores, utilizando apenas o NDVI, foi possível identificar quais as parcelas que efetivamente foram cultivadas e quais não foram.

Este procedimento foi possível através da análise dos gráficos NDVI para cada cultura, onde se estabeleceu um valor limite para cada cultura, valor esse que era o limiar para afirmar se existia ou não plantação naquela parcela. Seguidamente, foi feita a análise com as imagens, através da criação de um programa em Matlab®. A partir deste, se o valor limite estabelecido para cada cultura não estivesse presente em pelo menos 5 das 13 imagens da série temporal, era considerado que uma determinada parcela não apresentava esse tipo de

cultura. Depois de realizados estes passos foi possível identificar qual o número de parcelas declaradas erradamente como contendo uma dada cultura, avaliando assim o controlo físico efetuado na campanha de 2016.

2.3.2 Classificação de imagens

A classificação de imagem é o processo pelo qual os píxeis são agrupados em classes (previamente definidas) de acordo com as suas características espectrais. Para o estudo proposto para este tema foram utilizados três classificadores, sendo eles o de Máxima Verosimilhança, as Redes Neurais e o da Distância de Mahalanobis. O classificador de Máxima Verosimilhança pressupõe que as estatísticas de cada classe em cada banda estão distribuídas de forma equivalente e calcula a probabilidade de um determinado pixel pertencer a uma classe específica. Todos os píxeis são classificados a menos que seja atribuído um valor limite para a probabilidade. Cada píxel é atribuído à classe que tem maior probabilidade, ou seja, a máxima verosimilhança. Se a probabilidade mais alta for menor que um limite que tenha sido estabelecido, então o píxel permanece não classificado.

Por sua vez as Redes Neurais consistem numa técnica de classificação supervisionada avançada de redes neurais. É possível escolher um número de camadas ocultas a utilizar e a função a aplicar de entre logarítmica ou hiperbólica. Segundo Pádua [2000], as Redes Neurais são sistemas paralelos constituídos por unidades de processamento (nós) capazes de aplicar funções matemáticas simples aos dados recebidos. As funções matemáticas, inspiradas no sistema nervoso central, em particular o cérebro, são capazes de realizar variadas atividades assim como reconhecer padrões.

Já o classificador de Distância de Mahalanobis é um classificador de distância, sensível à direção, que usa estatísticas para cada classe. É semelhante ao classificador de Máxima Verosimilhança, mas assume que todas as covariâncias da classe são iguais. Todos os píxeis são classificados de acordo com a classe que se encontra mais próxima, a não ser que seja especificado um limite de distância, podendo existir píxeis que não são classificados se não cumprirem esse limite.

Para serem efetuados testes e retirar conclusões de quais os melhores resultados, foram consideradas diferentes combinações de bandas para a classificação das imagens: (a) a utilização de imagens com 4 bandas (VIS+NIR); (b) a utilização de imagens com 7 bandas (VIS+NIR+Red-Edge); (c) a utilização de imagens com 9 bandas (VIS+NIR+Red-Edge+SWIR); e ainda (d) a utilização de imagens resultantes da junção dos diversos índices gerados (Figura 2.5 a 2.8, respetivamente). No Anexo III estão presentes as imagens de cada índice para cada data.

Com base nos gráficos com a série temporal do NDVI resultantes da etapa anterior, e comparando as curvas com o valor médio do NDVI para cada uma das culturas, identificaram-se as datas ao longo do ciclo fenológico em que seria possível obter uma melhor discriminação entre os diferentes tipos de culturas.

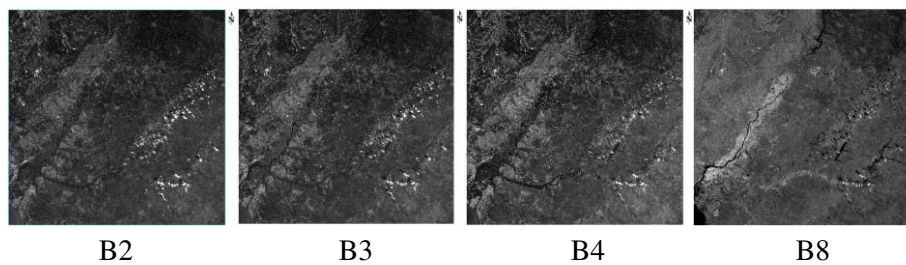


Figura 2.5: Combinação das 4 bandas para as três imagens utilizadas no estudo.

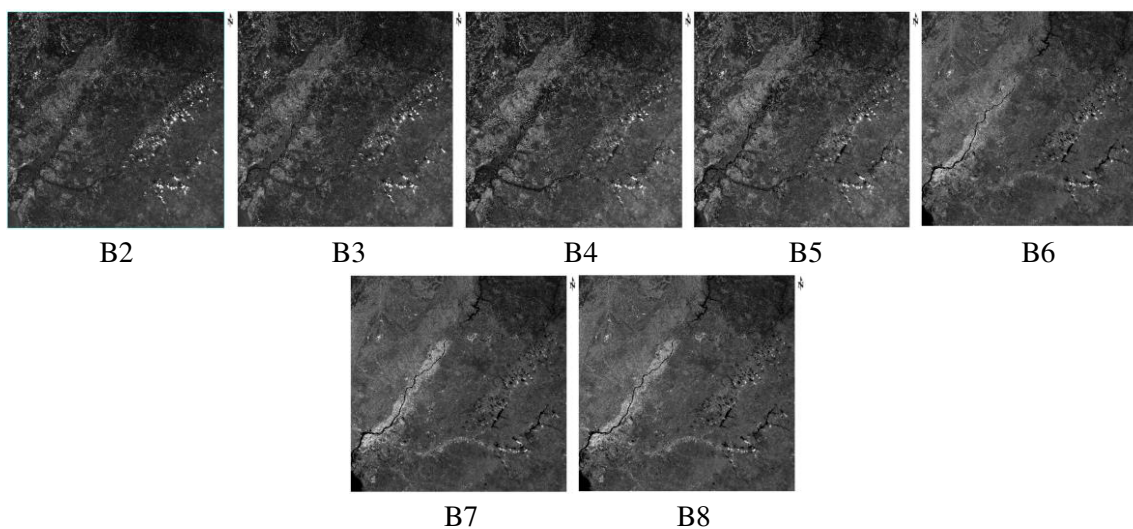


Figura 2.6: Combinação das 7 bandas para as três imagens utilizadas no estudo.

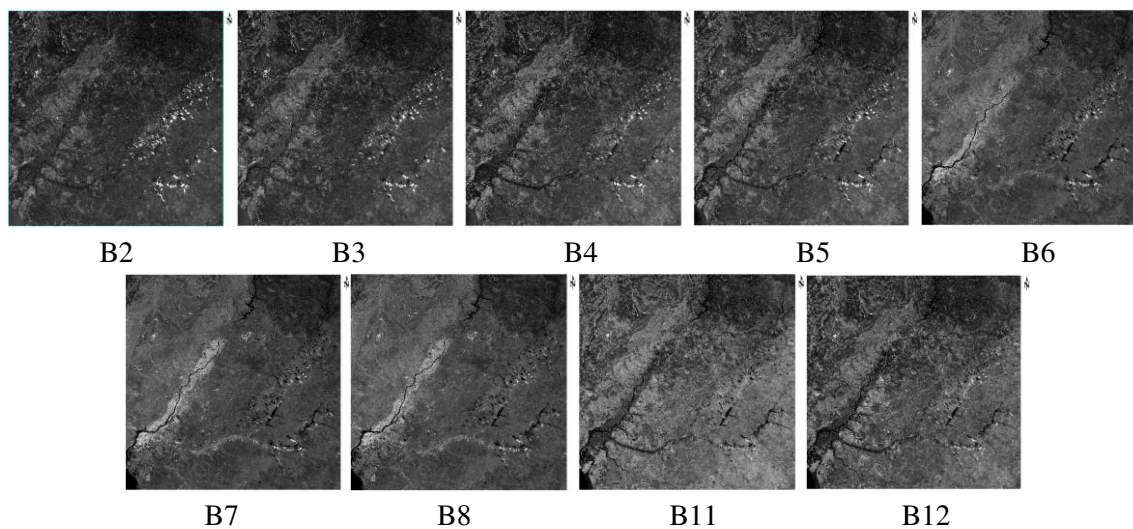


Figura 2.7: Combinações de 9 bandas para as três imagens utilizadas no estudo.

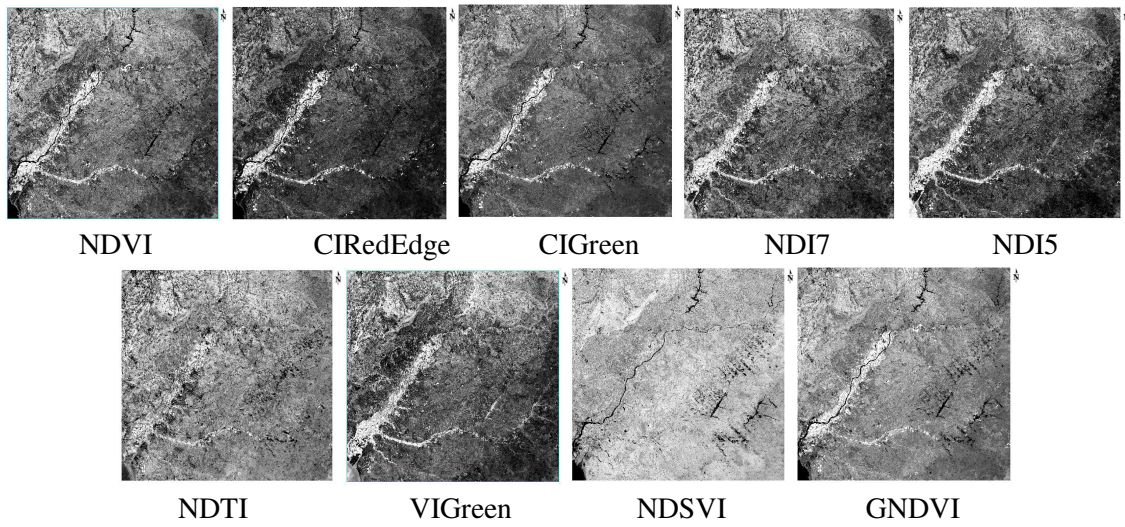


Figura 2.8: Combinação dos índices utilizados para o estudo.

3. Resultados

3.1 Índices de vegetação

As culturas utilizadas para o presente estudo foram 4, sendo elas pastagens, arroz, milho e tomate. As culturas foram ainda separadas consoante o seu crescimento, através da análise dos gráficos resultantes do índice de vegetação NDVI para cada cultura. Assim, foi possível separar as culturas de tomate em tomate precoce e tardio e milho em milho precoce e tardio. Também as pastagens foram separadas em pastagens arbustivas e permanentes com a análise do atributo referente ao tipo de cultura controlada, associado à tabela referente às parcelas, assim como o arroz em arroz e elemento linear de arroz, pelo mesmo método.

O início do ciclo vegetativo das culturas de irrigação é caracterizado por um baixo valor do NDVI com um aumento gradual ao longo do período de crescimento da planta e aumento da área foliar até atingir um máximo, geralmente superior a 0.7, designada fase de maturação, seguida uma fase de diminuição da atividade fotossintética a que corresponde a diminuição rápida do valor do NDVI até valores na ordem dos 0.3, associada à fase de senescência, ou seja, o envelhecimento da planta. O ciclo de crescimento e decréscimo do NDVI é caracterizado pela sua duração, valor máximo, amplitude e outros parâmetros característicos da curva, os quais permitem determinar, de forma inequívoca, se ocorreu plantação de uma cultura agrícola em alguma altura do ano, e por outro lado, com base na declaração do agricultor, determinar o nível de confiança dessa declaração.

Excetuando as pastagens, a cultura de milho é a mais representativa na área de estudo, sendo cultivado sobretudo em condições de regadio que caracterizam os ambientes mediterrâneos. Por norma o milho é semeado nos meses de março/maio e colhido nos meses de agosto/setembro. Na Figura 3.1 é possível observar a evolução das parcelas de milho tardio, sobrepostas à imagem NDVI, em quatro datas da série temporal, onde se verifica que no mês de junho a zona é mais escura, correspondendo a valores mais baixos, o que indica a presença de solo nu ou sem vegetação e no mês de agosto verificamos valores próximos de 1 que indicam a presença de vegetação densa e vigorosa.

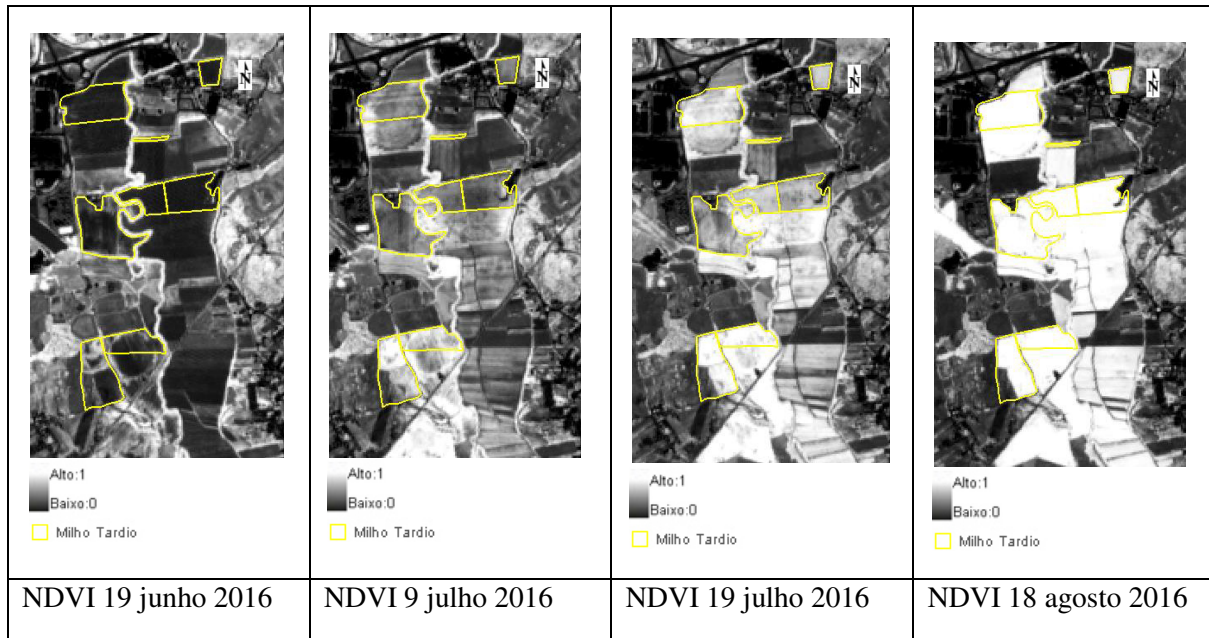


Figura 3.1: Imagens NDVI para zona de cultura de Milho Tardio.

Na Figura 3.2 e Figura 3.3 encontram-se representados os gráficos de NDVI para as diferentes parcelas das culturas de milho, tomate, arroz e pastagens. Analisando os gráficos presentes nas figuras é possível verificar que em todos existe uma pequena quebra das linhas perto do DOY (*Day Of the Year*) 190, que corresponde ao dia 9 de julho. Esta pequena diminuição deve-se ao facto de no ano de 2016 as temperaturas terem sido muito elevadas nesta altura do ano, que corresponde ao mês de julho, sendo que possivelmente as regas foram ajustadas, verificando-se um novo crescimento da curva para os valores esperados. Nas culturas de arroz e pastagens esta quebra não é tão evidente porque o arroz se encontra num meio húmido, sendo a sua resposta espectral a mesma, enquanto que as pastagens são já bastantes secas, não sofrendo grande influência pelo aumento repentino da temperatura.

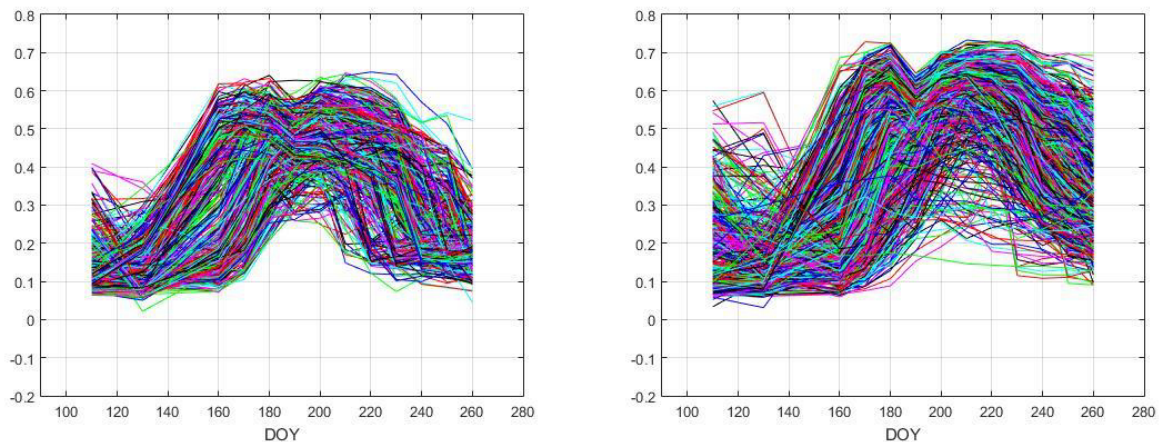


Figura 3.2: NDVI Milho (esquerda) e Tomate (direita).

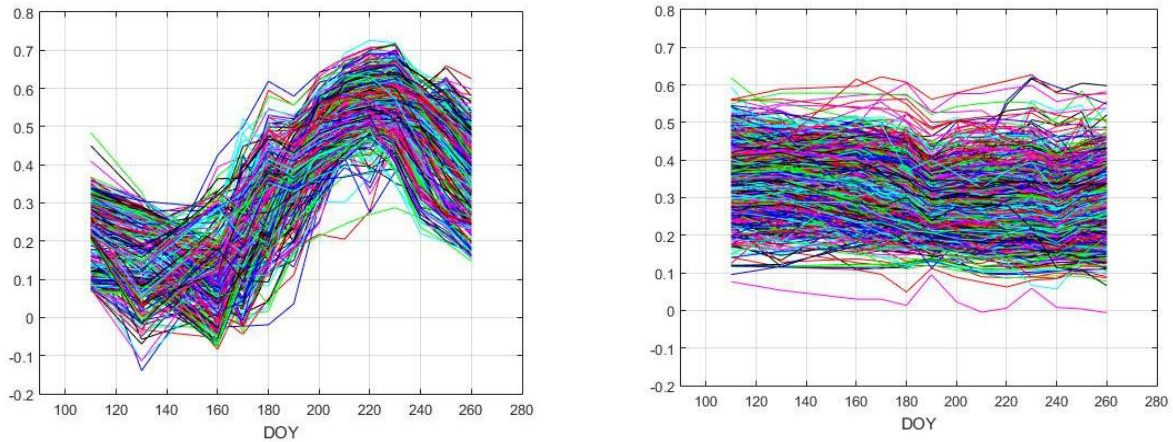


Figura 3.3: NDVI Arroz (esquerda) e Pastagens (direita).

É possível verificar diferenças nas curvas para uma mesma cultura, estando o seu comportamento fenológico desfasado no tempo, nomeadamente para as culturas de milho e de tomate, havendo parcelas que foram plantadas mais cedo, ou seja, atingem o valor máximo de NDVI primeiro (precoces), e outras que foram plantadas mais tardiamente, atingindo o valor máximo de NDVI depois (tardias). Em relação às pastagens e ao arroz, estas foram separadas em arbustivas e permanentes e em arroz e elemento linear de arroz, através do atributo existente na tabela agregada ao ficheiro vetorial das parcelas.

Da Figura 3.4 à Figura 3.7 estão presentes os gráficos com as curvas do NDVI para as culturas devidamente separadas e tratadas, onde é possível verificar a diferença de comportamento existente dentro da mesma cultura. Este processamento foi possível devido à análise feita a cada parcela para cada cultura, através do seu valor NDVI e também do valor do DOY em que esta começava a ter senescência.

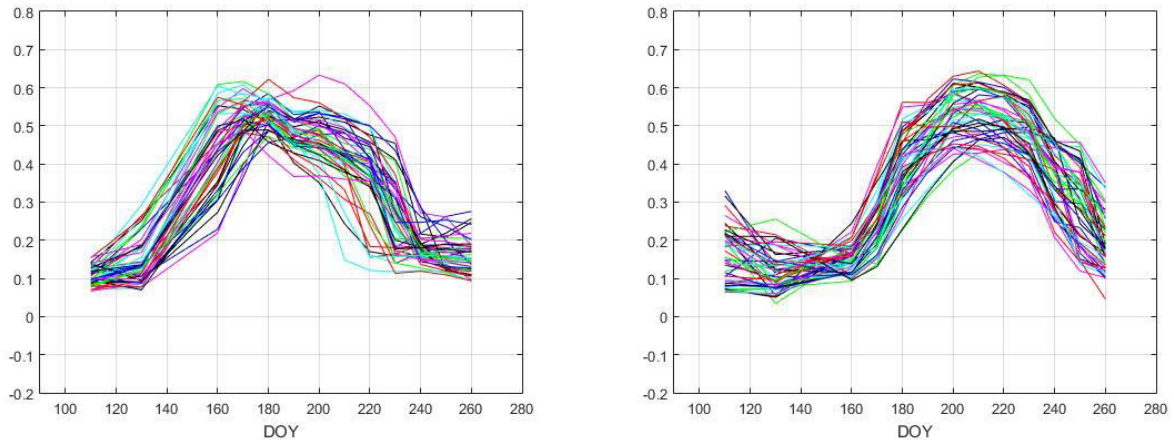


Figura 3.4: NDVI Tomate Precoce (esquerda) e Tomate Tardio (direita).

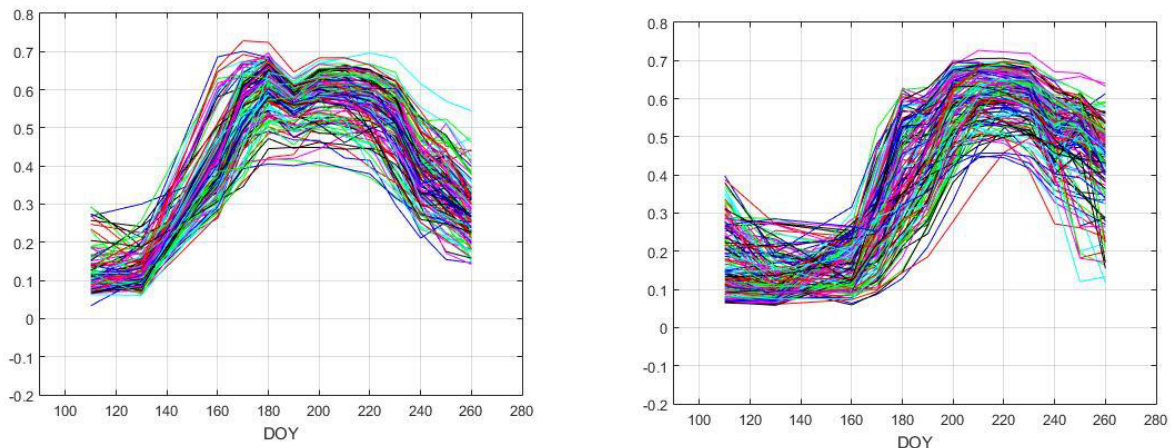


Figura 3.5: NDVI Milho Precoce (esquerda) e Milho Tardio (direita).

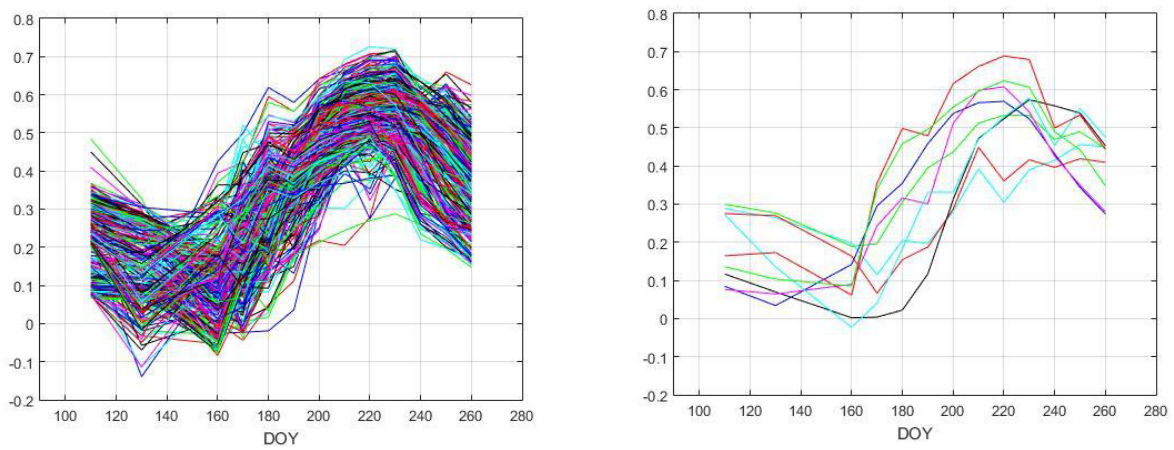


Figura 3.6: NDVI Arroz (esquerda) e Elemento Linear de Arroz (direita).

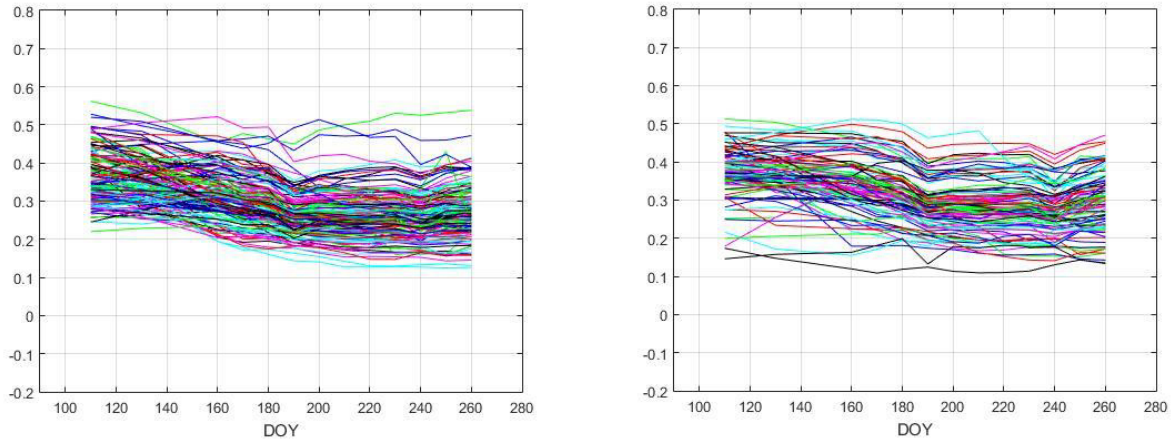


Figura 3.7: NDVI Pastagens Permanentes (esquerda) e Pastagens Arbustivas (direita).

Também foram testados outros índices de vegetação, sendo que graficamente não foram retiradas conclusões significativas relativamente a cada cultura. De salientar que para a cultura de pastagens, não existiam alterações nos gráficos com a aplicação dos diferentes índices, sendo que estes não são apresentados nos resultados finais. No entanto, os índices que são afetos pelas bandas do *Red-Edge*, como é o caso do CIRedEdge, têm a particularidade de serem mais estáveis no início e no fim do ciclo, enquanto que na fase de maturação sofrem picos relevantes.

Na Figura 3.8 estão presentes os gráficos resultantes deste índice para três das culturas em estudo.

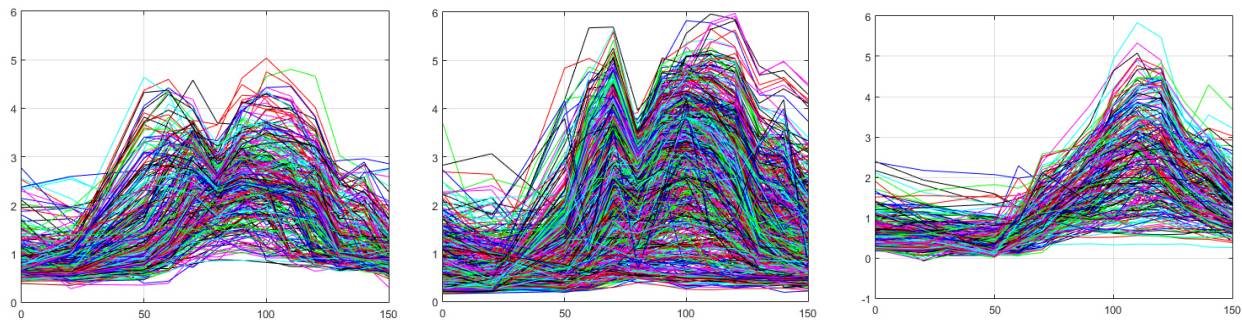


Figura 3.8: CIRedEdge para a cultura de tomate (esquerda), milho (centro) e arroz (direita).

Ao observar os gráficos presentes nas figuras anteriores é possível verificar que existe uma maior desorganização das curvas de cada parcela ao longo da série temporal para a cultura de arroz, em comparação com as curvas no gráfico do índice NDVI (Figuras 3.2 e 3.3). Este acontecimento deve-se ao facto de as bandas do *Red-Edge*, presentes no índice CIRedEdge, serem mais sensíveis à mudança da refletância da vegetação nas bandas do infravermelho próximo, o que não se verifica no arroz uma vez que este contém água na sua plantação.

Também para o índice GNDVI foram obtidos gráficos com algumas diferenças entre as culturas, comparativamente com os gráficos do NDVI. Na Figura 3.9 estão presentes os gráficos para as três culturas também anteriormente apresentadas.

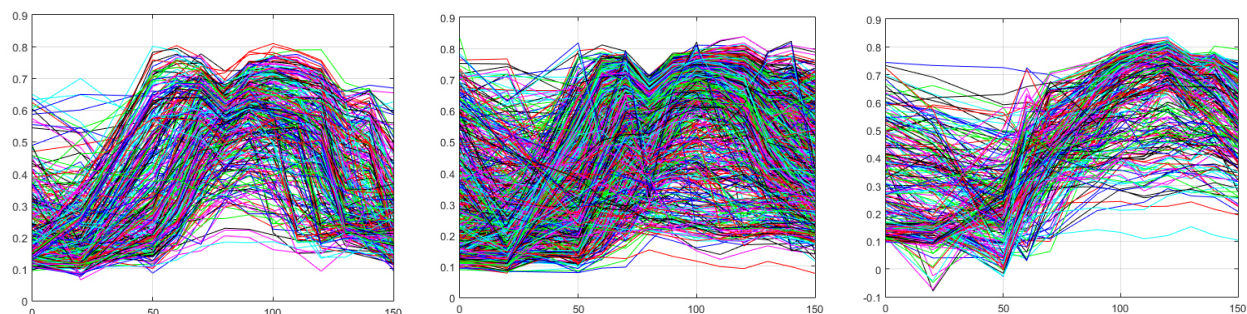


Figura 3.9: GNDVI para a cultura de tomate (esquerda), milho (centro) e arroz (direita).

Sendo este índice afeto por bandas na zona do verde e do infravermelho próximo, o seu comportamento deveria ser semelhante ao NDVI, onde entram as bandas do infravermelho próximo e do vermelho. No entanto, apesar de se encontrarem semelhanças, é possível verificar que existem parcelas que têm um comportamento diferente do esperado dentro de cada cultura, sendo que os picos existentes estão muito mais evidenciados do que nos gráficos do NDVI. Neste caso, apenas o tomate poderia ser considerado como a cultura com o comportamento mais semelhante. Uma justificação possível é o facto de este ter uma resposta espectral semelhante na banda do vermelho e do verde, sendo que o infravermelho próximo é comum a ambos os índices.

Com os valores de NDVI obtidos para cada cultura, é possível afirmar que a parcela teve ocupação cultural ou não teve ocupação cultural, não havendo possibilidade de distinguir se a parcela declarada é milho ou tomate, por exemplo. Com a análise dos respetivos gráficos, foi possível dizer qual o número de parcelas que efetivamente tem algum tipo de plantação e qual o número de parcelas que não teve ocupação cultural, apesar de ter sido controlada. Na Tabela 3.1 é visível o número de parcelas declaradas e controladas fisicamente para cada cultura e o número de parcelas que tiveram ocupação cultural.

Tabela 3.1: Controlo das parcelas controladas fisicamente.

Cultura	Nº Parcelas Controladas	Nº Parcelas Com Plantação	Diferença entre Parcelas Controladas e Parcelas Com Plantação	Valor NDVI
Arroz	306	303	3	0.3
Milho	675	627	48	0.4
Tomate	400	389	11	0.3
Pastagens	915	815	100	0.2

A análise da série temporal do NDVI de cada parcela declarada permitiu identificar as parcelas em que não foi registada qualquer atividade agrícola (quarta coluna da tabela 3.1). Para as restantes parcelas procedeu-se à classificação da ocupação do solo com o propósito de validar a declaração do agricultor quanto à cultura declarada.

3.2 Classificação de imagens

Com base no gráfico da evolução temporal do NDVI (Figura 3.10) para cada uma das culturas agrícolas foram selecionadas 4 datas para a realização da classificação do uso do solo. A seleção destas 4 datas foi efetuada através da análise das curvas com o ciclo fenológico das culturas, identificando quais as datas que permitiam a maior separabilidade possível entre as diferentes classes consideradas.

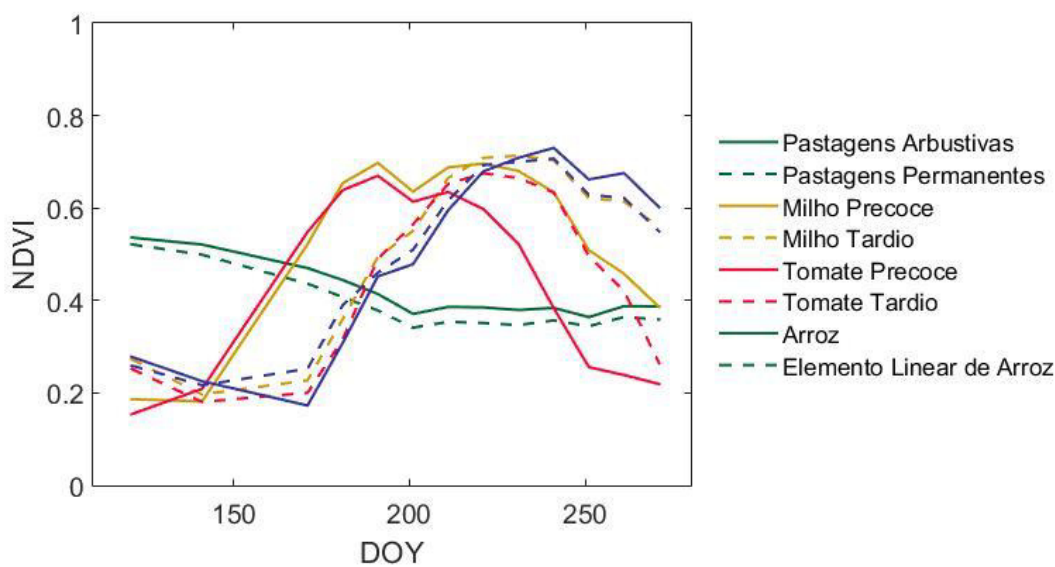


Figura 3.10: Gráfico com a média para cada cultura ao longo da série temporal.

De acordo com o gráfico da Figura 3.10, a classificação da imagem do dia 171 (19 de junho) permitirá fazer uma boa diferenciação entre as culturas de milho precoce e tardio, assim como entre tomate precoce e tardio. O arroz também tem um comportamento diferente das culturas de milho precoce e tomate precoce, mas semelhante ao milho tardio e ao tomate tardio, podendo existir confusão entre estas culturas. As culturas de pastagens arbustivas e pastagens permanentes têm um comportamento semelhante ao longo de toda a série temporal sendo as culturas que melhor se distinguirão das restantes, podendo existir confusão apenas próximo do DOY 181 (29 de julho), onde todas as culturas se cruzam, exceto o milho precoce e o tomate precoce.

As melhores datas para a possível diferenciação entre todo o tipo de culturas são as datas que correspondem ao DOY 171 (19 de junho), 191 (9 de julho), 201 (19 de julho) e 231 (18 de agosto). Também se verifica que o comportamento das pastagens arbustivas e pastagens permanentes é o mesmo, assim como do elemento linear de arroz e arroz, pelo que se optou por agregar estas sub-classes, gerando as classes pastagens e arroz, utilizadas como áreas de treino e validação.

Assim sendo, as áreas de treino são um conjunto de parcelas selecionadas de entre o conjunto total de parcelas correspondente a cada cultura, efetuando o mesmo processo para as áreas de validação. As áreas são utilizadas para a criação de classes ou categorias, necessárias para a classificação de imagem. Estas zonas foram escolhidas aleatoriamente, fazendo a divisão de metade do ficheiro vetorial de cada cultura para

validação e a outra metade para treino, tendo sempre o cuidado de verificar que a zona era coberta na totalidade em ambas os casos.

As áreas de treino e validação foram utilizadas de três formas diferentes. Primeiro começou por se utilizar 8 classes, de seguida 6 classes, e por fim apenas 4 classes, de forma a verificar quais os melhores resultados obtidos. A separação foi feita mediante a separação efetuada para as culturas, utilizando milho precoce, milho tardio, tomate precoce, tomate tardio, pastagens e arroz para 6 classes. Relativamente às 8 classes fez-se a diferenciação entre pastagens arbustivas e pastagens permanentes e também entre arroz e elemento linear de arroz. Para 4 classes apenas se diferenciou tomate, milho, arroz e pastagens. Devido ao comportamento das pastagens arbustivas e permanentes e do arroz e elemento linear de arroz, ao utilizar 8 classes é de esperar que os resultados não sejam tão bons como com 6 ou 4 classes.

A classificação das 4 imagens foi realizada com 3 classificadores no sentido de se avaliar qual o classificador que melhor exatidão tem para este problema concreto. Foram testados dois classificadores paramétricos Bayseanos, Máxima Verosimilhança e Distância de Mahalanobis e um classificador que usa aprendizagem máquina (“*machine learning*”), as Redes Neurais.

3.2.1 Classificador de Máxima Verosimilhança

Este método classifica os pixéis de acordo com a regra de Bayes, como referido anteriormente, que determina a probabilidade de um determinado pixel pertencer a uma determinada classe, através da sua resposta espectral e com base na sua localização, seguindo a distribuição normal gaussiana [Richards e Jia, 2006]. Essa probabilidade é dada pela equação 3.1:

$$p(w_i|x) = p(x|w_i) p(w_i) / p(x) \quad (3.1)$$

onde $p(w_i)$ é a probabilidade que cada classe w_i tem de ocorrer na imagem e x representa o valor do pixel em cada banda.

Às 4 datas escolhidas para a validação foram aplicados os três conjuntos de áreas de treino e validação, já referidos anteriormente, avaliando a sua exatidão global e os resultados da matriz de confusão de modo a determinar qual o melhor conjunto, verificando-se em todos os casos que é apenas o de quatro classes (Tabela 3.2), conjunto este que não gera confusão entre as mesmas.

Tabela 3.2: Resultados da exatidão global para as 4 datas e imagem de 4 bandas utilizando a Máxima Verosimilhança.

Classes	Exatidão global (%)			
	19 junho 2016	9 julho 2016	19 julho 2016	8 agosto 2016
8 classes	45.89%	53.11%	52.04%	47.14%
6 classes	60.93%	66.55%	67.26%	60.93%
4 classes	73.39%	86.58%	86.70%	74.73%

Na Tabela 3.3 é possível verificar os resultados obtidos para a imagem que corresponde a 9 bandas, B2-B3-B4-B5-B6-B7-B8-B11-B12, e para a imagem com todos os índices de vegetação, de 19 julho 2016, onde se verifica que a exatidão global apresenta um bom valor, cerca de 90% em ambos os casos, assim como a

matriz de confusão mostra que não existe grande confusão entre classes diferentes e os erros de omissão e comissão apresentam valores relativamente baixos.

Tabela 3.3: Resultados estatísticos para as imagens das 9 bandas e dos índices de 19 julho 2016.

	Bandas				Índices			
Exatidão Global	89.72%				87.99%			
Coefficiente K	85.59%				83.13%			
Confusão (%)								
	Milho		Pastagens		Arroz		Tomate	
	Bandas	Índices	Bandas	Índices	Bandas	Índices	Bandas	Índices
Milho	88.25	84.54	2.38	2.13	6.93	10.65	25.88	13.26
Pastagens	7.23	8.25	93.65	95.23	5.40	6.78	41.24	1.80
Arroz	3.74	2.29	3.91	2.57	87.13	80.74	17.56	1.27
Tomate	0.78	4.92	0.07	0.07	0.54	1.83	85.52	83.66
Precisões (%)								
	Precisão Produtor		Precisão Utilizador		Erro Comissão		Erro Omissão	
	Bandas	Índices	Bandas	Índices	Bandas	Índices	Bandas	Índices
Milho	88.25	88.25	87.57	87.57	12.43	12.43	11.75	11.75
Pastagens	93.65	93.65	90.45	90.45	9.55	9.55	6.35	6.35
Arroz	87.13	97.13	84.08	84.08	15.92	15.92	12.87	12.87
Tomate	85.52	85.52	97.91	97.91	2.09	2.09	14.48	14.48

Sendo este classificador o mais rápido na obtenção de resultados, foi com este que se fizeram os testes para todas as possibilidades de conjugações de bandas em imagens, nas quatro datas escolhidas, testando posteriormente os outros dois classificadores na imagem das bandas que melhores resultados obteve e também para a imagem dos índices.

Na Tabela 3.4 são apresentados os valores da exatidão global obtidos para todas as imagens testadas, nas quatro datas escolhidas para a classificação, utilizando 4 classes, sendo este o conjunto que obteve melhores resultados.

Tabela 3.4: Exatidão Global para as 4 datas escolhidas para a validação para os 4 tipos de imagens testadas.

Imagens	B2_B3_B4_B8	B2_B3_B4_B5_B6_B7_B8	B2_B3_B4_B5_B6_B7_B8_B11_B12	Índices
19 junho 2016	73.39%	79.58%	78.97%	78.04%
9 julho 2016	86.58%	89.13%	87.99%	85.70%
19 julho 2016	86.70%	89.33%	89.73%	87.99%
18 agosto 2016	74.73%	79.24%	85.17%	83.62%

Ao observar a tabela é notório que os melhores resultados estão todos no dia 19 julho 2016, tanto em termos de bandas como em termos de índices. Verificamos que a utilização das bandas resulta sempre numa maior exatidão global quando comparado com a utilização de índices.

Como se constrói uma imagem multibanda com todas as 9 bandas (B2-B3-B4-B5-B6-B7-B8-B11-B12) e estas não são combinadas entre si, contrariamente ao que acontece quando se constrói a imagem para os índices, onde existem cálculos com bandas para cada índice de vegetação, visível na Tabela 2.1, presente no subcapítulo 2.3.1, os resultados para a exatidão global com as 9 bandas são melhores do que os resultados com os índices. No Anexo I encontram-se as imagens resultantes para cada data.

3.2.2 Classificador de Redes Neurais

O classificador de Redes Neurais consiste num conjunto de redes neuronais artificiais baseadas na estrutura do cérebro. Existe uma série de funções que podem ser utilizadas sendo a mais comum a função logarítmica. Esta função, presente na equação 3.2 é dada por:

$$f(x_i) = 1 / (1 + e^{x_i}) \quad (3.2)$$

onde x_i representa os valores de cada classe e $f(x)$ o valor da função resultante do classificador que origina um conjunto de valores de saída. De salientar que o conjunto dos valores iniciais têm pesos associados, sendo que para cada valor x_i existe um peso w_i .

Para a data de 19 julho de 2016, é possível verificar que a exatidão global obtida com as bandas e com os índices, apresenta valores muito idênticos, com uma diferença de apenas duas centésimas, visível na Tabela 3.5.

Tabela 3.5: Resultados estatísticos para as imagens das 9 bandas e dos índices de 19 julho 2016.

	Bandas				Índices			
Exatidão Global	87.17%				87.19%			
Coefficiente K	81.92%				82.02%			
Confusão (%)								
	Milho		Pastagens		Arroz		Tomate	
	Bandas	Índices	Bandas	Índices	Bandas	Índices	Bandas	Índices
Milho	92.39	87.89	5.69	5.29	21.70	14.42	11.32	10.35
Pastagens	5.69	6.15	92.27	92.36	6.69	7.11	5.18	3.12
Arroz	0.96	0.55	1.99	1.68	71.27	74.89	0.23	0.22
Tomate	0.97	5.41	0.04	0.68	0.34	3.58	83.28	86.31
Precisões (%)								
	Precisão Produtor		Precisão Utilizador		Erro Comissão		Erro Omissão	
	Bandas	Índices	Bandas	Índices	Bandas	Índices	Bandas	Índices
Milho	92.39	87.89	74.95	78.00	25.05	22.00	7.61	12.11
Pastagens	92.27	92.36	91.30	91.70	8.70	8.30	7.73	7.64
Arroz	71.27	74.89	91.80	93.75	8.20	6.25	28.73	25.11
Tomate	83.28	86.31	97.83	86.97	2.17	13.03	16.72	13.69

Em termos de exatidão global, este classificador mostrou ser inferior ao de Máxima Verosimilhança. No entanto, quando avaliamos a confusão entre classes percebemos que neste classificador os valores são inferiores, apresentando uma menor confusão entre classes diferentes, como é o caso das pastagens, que quando avaliadas em relação ao milho apresentam valores de aproximadamente 6% para as bandas e para os índices, enquanto que no classificador anterior esses valores eram superiores a 7% e 8%, respetivamente para as bandas e para os índices.

Relativamente às 8 e 6 classes e fazendo a comparação com o classificador anterior, retiramos a mesma conclusão, sendo que os melhores resultados obtidos são com o classificador de Redes Neurais. A Tabela 3.6 mostra a comparação entre os dois classificadores para 8 classes e a Tabela 3.7 para 6, utilizando a imagem de 4 bandas (B2-B3-B4-B8) e a imagem dos índices.

Tabela 3.6: Comparação de resultados estatísticos com 8 classes para as imagens das bandas e dos índices de 19 julho 2016.

	Máxima Verosimilhança (MV)								Redes Neurais (NN)							
Exatidão Global	52.04%								66.59%							
Coefficiente K	42.92%								57.68%							
Confusão (%)																
	Milho Precoce		Milho Tardio		Arroz		Elemento Linear de Arroz		Tomate Precoce		Tomate Tardio		Pastagens Arbustivas		Pastagens Permanentes	
	MV	RN	MV	RN	MV	RN	MV	RN	MV	RN	MV	RN	MV	RN	MV	RN
Milho Precoce	78.10	76.96	39.58	30.45	3.06	1.75	3.57	4.33	0.85	1.44	3.61	3.34	0.48	0.85	0.11	0.33
Milho Tardio	4.34	7.07	24.15	35.89	4.99	18.30	5.70	26.06	0.62	0.26	4.01	0.86	4.74	4.99	2.64	1.58
Arroz	0.09	0.24	3.25	0.83	18.14	74.32	37.51	49.10	0.03	0.01	0.17	0.70	0.60	1.67	0.57	0.28
El. Linear Arroz	0.23	0.00	2.05	0.00	64.97	0.00	32.64	0.00	0.97	0.00	0.22	0.00	0.13	0.00	0.01	0.00
Tomate Precoce	0.39	0.69	0.16	0.23	0.20	0.25	0.25	0.12	53.07	67.06	11.93	42.01	0.05	0.80	0.01	0.08
Tomate Tardio	2.15	4.53	1.32	5.33	0.73	2.32	1.70	4.16	13.00	29.38	44.55	47.12	0.15	0.69	0.04	0.17
P. Arbustivas	3.63	0.04	13.12	0.03	0.82	0.00	4.47	0.00	2.91	0.01	5.63	0.04	11.29	0.14	6.20	0.03
P. Permanentes	4.99	10.47	9.87	27.24	1.12	3.06	7.45	16.23	0.37	1.83	0.78	5.94	61.76	90.86	82.47	97.54
Precisões (%)																
	Precisão Produtor		Precisão Utilizador				Erro Comissão				Erro Omissão					
	MV	RN	MV	RN	MV	RN	MV	RN	MV	RN	MV	RN				
Milho Precoce	78.10	76.96	57.82	62.88	42.18	37.12	21.90	23.04								
Milho Tardio	24.15	35.89	53.75	50.00	46.25	50.00	75.85	64.11								
Arroz	18.14	74.32	18.14	87.15	37.98	12.85	81.86	25.68								
El. Linear Arroz	32.64	0.00	7.35	0.00	92.65	0.00	67.36	100.00								
Tomate Precoce	53.07	67.06	83.64	65.65	16.36	34.35	46.93	32.94								
Tomate Tardio	44.55	47.12	64.39	43.37	35.61	56.63	55.45	52.88								
P. Arbustivas	11.29	0.14	13.89	31.43	86.11	68.57	88.71	99.86								
P. Permanentes	82.47	97.54	79.19	70.36	20.81	29.64	17.53	2.46								

Analisando os resultados presentes na tabela anterior vemos que a exatidão global é mais elevada no classificador de Redes Neurais do que no de Máxima Verosimilhança. Relativamente à confusão entre classes, dependendo da classe, tanto um classificador como outro têm melhores resultados. Com 8 classes existe uma grande confusão entre as mesmas, principalmente entre classes com comportamento semelhante, como no caso das pastagens arbustivas e permanentes e do arroz e elemento linear de arroz, como referido anteriormente após análise do gráfico da média das médias do NDVI.

Tabela 3.7: Comparação de resultados estatísticos com 6 classes para as imagens das bandas e dos índices de 19 julho 2016.

	Máxima Verosimilhança (MV)						Redes Neurais (RN)					
Exatidão Global	67.26%						63.86%					
Coefficiente K	58.67%						51.06%					
Confusão (%)												
	Milho Precoce		Milho Tardio		Arroz		Tomate Precoce		Tomate Tardio		Pastagens	
	MV	RN	MV	RN	MV	RN	MV	RN	MV	RN	MV	RN
Milho Precoce	78.11	72.31	39.90	32.66	3.39	18.60	0.88	0.77	3.63	1.68	0.19	0.23
Milho Tardio	4.40	2.52	26.43	10.93	5.47	20.31	0.71	0.10	4.39	0.34	3.42	0.24
Arroz	0.20	0.10	4.23	0.15	76.15	48.77	0.41	0.00	0.25	0.07	0.48	0.82
Tomate Precoce	0.40	0.04	0.16	0.00	0.25	0.01	53.19	9.33	11.97	9.98	0.02	0.04
Tomate Tardio	2.15	6.91	1.34	3.89	0.89	1.84	13.02	86.94	44.58	73.55	0.06	0.36
Pastagens	6.37	18.12	17.31	52.37	2.43	10.47	0.85	2.86	1.53	14.38	82.73	98.31
Precisões (%)												
	Precisão Produtor		Precisão Utilizador		Erro Comissão		Erro Omissão					
	MV	RN	MV	RN	MV	RN	MV	RN				
Milho Precoce	78.11	72.31	57.48	50.83	42.52	49.17	21.89	27.69				
Milho Tardio	26.43	10.93	53.93	28.63	46.07	71.37	73.57	89.07				
Arroz	76.15	48.77	93.67	95.74	6.33	4.26	23.85	51.23				
Tomate Precoce	53.19	9.33	83.52	53.87	16.48	46.13	46.81	90.67				
Tomate Tardio	44.58	73.55	64.32	35.21	35.68	64.79	55.42	26.45				
Pastagens	82.73	98.31	89.65	75.54	10.35	24.46	17.27	1.69				

Juntando as classes de arroz e elemento linear de arroz numa só, assim como as pastagens arbustivas e permanentes, a exatidão global do classificador de Redes Neurais volta novamente a ser inferior ao valor do classificador de Máxima Verosimilhança, como se constatou aquando da utilização de apenas 4 classes. Neste caso e na generalidade dos resultados, verifica-se que os resultados de confusão entre classes são inferiores quando se utiliza as Redes Neurais. No Anexo II estão presentes as imagens resultantes.

3.2.3 Classificador de Distância de Mahalanobis

Relativamente a este classificador, apenas foram testados os conjuntos de 8 e 6 classes para a imagem de 4 bandas (B2-B3-B4-B8), sendo que os resultados são semelhantes aos obtidos para o classificador de Máxima Verosimilhança, como se observa nas Tabelas 3.8 e 3.9.

Tabela 3.8: Comparação de resultados estatísticos com 8 classes para as imagens das bandas e dos índices de 19 julho 2016.

	Máxima Verosimilhança (MV)								Distância de Mahalanobis (DM)							
Exatidão Global	52.04%								51.40%							
Coefficiente K	42.92%								42.09%							
Confusão (%)																
	Milho Precoce		Milho Tardio		Arroz		Elemento Linear de Arroz		Tomate Precoce		Tomate Tardio		Pastagens Arbustivas		Pastagens Permanentes	
	MV	DM	MV	DM	MV	DM	MV	DM	MV	DM	MV	DM	MV	DM	MV	DM
Milho Precoce	78.10	74.80	39.58	37.30	3.06	0.99	3.57	0.27	0.85	2.50	3.61	15.38	0.48	0.99	0.11	0.08
Milho Tardio	4.34	8.16	24.15	21.68	4.99	7.91	5.70	8.23	0.62	0.28	4.01	2.16	4.74	10.67	2.64	4.77
Arroz	0.09	0.21	3.25	4.87	18.14	26.83	37.51	39.27	0.03	0.02	0.17	0.23	0.60	9.93	0.57	4.82
El. Linear Arroz	0.23	0.67	2.05	7.91	64.97	59.37	32.64	29.46	0.97	0.45	0.22	0.48	0.13	3.41	0.01	0.93
Tomate Precoce	0.39	2.18	0.16	1.37	0.20	1.05	0.25	6.09	53.07	76.33	11.93	38.22	0.05	0.32	0.01	0.08
Tomate Tardio	2.15	1.68	1.32	1.91	0.73	1.88	1.70	3.18	13.00	14.19	44.55	33.19	0.15	4.49	0.04	0.94
P. Arbustivas	3.63	8.56	13.12	20.47	0.82	1.17	4.47	7.53	2.91	6.02	5.63	9.79	11.29	19.39	6.20	15.57
P. Permanentes	4.99	3.74	9.87	4.51	1.12	0.79	7.45	5.97	0.37	0.20	0.78	0.54	61.76	50.80	82.47	72.81
Precisões (%)																
	Precisão Produtor		Precisão Utilizador		Erro Comissão				Erro Omissão							
	MV	DM	MV	DM	MV	DM	MV	DM	MV	DM	MV	DM	MV	DM	MV	DM
Milho Precoce	78.10	74.80	57.82	55.35	42.18	44.65	21.90	25.20								
Milho Tardio	24.15	21.68	53.75	38.87	46.25	61.13	75.85	78.32								
Arroz	18.14	26.83	18.14	49.50	37.98	50.50	81.86	73.17								
El. Linear Arroz	32.64	29.46	7.35	6.33	92.65	93.67	67.36	70.54								
Tomate Precoce	53.07	76.33	83.64	66.89	16.36	33.11	46.93	23.67								
Tomate Tardio	44.55	33.19	64.39	47.12	35.61	52.88	55.45	66.81								
P. Arbustivas	11.29	19.39	13.89	12.00	86.11	88.00	88.71	80.61								
P. Permanentes	82.47	72.81	79.19	81.99	20.81	18.01	17.53	27.19								

Sendo os resultados em ambos os classificadores semelhantes, optou-se por testar apenas mais um conjunto de classes para a imagem de 19 julho 2016 e verificar qual o que iria ser utilizado para comparar com o classificador de Redes Neurais, sendo que foi o de Máxima Verosimilhança, como referido anteriormente, por se ter verificado mais rápido em termos de processamento de dados.

Tabela 3.9: Comparação de resultados estatísticos com 6 classes para as imagens das bandas e dos índices de 19 julho 2016.

	Máxima Verosimilhança (MV)						Distância de Mahalanobis (DM)					
Exatidão Global	67.26%						69.47%					
Coefficiente K	58.67%						60.30%					
Confusão (%)												
	Milho Precoce		Milho Tardio		Arroz		Tomate Precoce		Tomate Tardio		Pastagens	
	MV	DM	MV	DM	MV	DM	MV	DM	MV	DM	MV	DM
Milho Precoce	78.11	74.83	39.90	37.31	3.39	0.89	0.88	2.50	3.63	15.39	0.19	0.25
Milho Tardio	4.40	8.77	26.43	23.75	5.47	9.42	0.71	0.32	4.39	2.33	3.42	6.85
Arroz	0.20	0.38	4.23	9.99	76.15	79.27	0.41	0.09	0.25	0.37	0.48	6.88
Tomate Precoce	0.40	3.26	0.16	4.89	0.25	4.91	53.19	81.21	11.97	43.80	0.02	0.49
Tomate Tardio	2.15	2.05	1.34	3.28	0.89	2.46	13.02	14.79	44.58	35.49	0.06	1.85
Pastagens	6.37	10.71	17.31	20.78	2.43	3.04	0.85	1.09	1.53	2.63	82.73	83.67
Precisões (%)												
	Precisão Produtor		Precisão Utilizador		Erro Comissão		Erro Omissão					
	MV	DM	MV	DM	MV	DM	MV	DM				
Milho Precoce	78.11	74.83	57.48	55.35	42.52	44.65	21.89	25.17				
Milho Tardio	26.43	23.75	53.93	37.78	46.07	62.22	73.57	76.25				
Arroz	76.15	79.27	93.67	76.01	6.33	23.99	23.85	20.73				
Tomate Precoce	53.19	81.21	83.52	59.47	16.48	40.53	46.81	18.79				
Tomate Tardio	44.58	35.49	64.32	45.09	35.68	54.91	55.42	64.51				
Pastagens	82.73	83.67	89.65	86.93	10.35	13.07	17.27	16.33				

3.3 Validação de declarações

O ponto de partida para a validação das declarações dos agricultores é a análise feita às parcelas de cada cultura. Com o comportamento das parcelas e os valores de NDVI obtidos é possível validar as declarações, efetuando a divisão entre parcelas corretamente declaradas e parcelas que suscitam dúvidas por não corresponderem ao comportamento esperado.

As parcelas que suscitam dúvidas podem surgir devido à imagem da série temporal analisada ter nuvens ou fumo impercetível, não existir vegetação na zona da parcela na altura em que foram adquiridas as imagens, os agricultores declararem erradamente a cultura existente na parcela ou ainda a utilização da mesma para outro tipo de culturas antes da época da cultura declarada, existindo assim um comportamento diferente do esperado.

Finalmente, faz-se uma análise global de todos os dados obtidos, chegando à conclusão que a parcela tem que ter um valor de NDVI superior a 0.4 para se considerar que existiu plantação no caso do milho, superior a 0.3 no caso do arroz e do tomate e 0.2 no caso das pastagens, e a imagem tem que ter uma exatidão global acima dos 80% quando classificada utilizando o classificador de Redes Neurais para poder ser considerada como apta para avaliação. Posteriormente é feita a análise visual e a análise dos resultados estatísticos da matriz de confusão, onde não pode existir confusão entre classes diferentes, ou seja, as percentagens têm que se encontrar com valores baixos, excetuando as que fazem correspondência com a mesma classe.

Na Tabela 3.10 estão representados os valores obtidos para cada cultura para a imagem de 19 julho 2016, com o classificador de Redes Neurais.

Tabela 3.10: Percentagem de parcelas validadas e parcelas suspeitas para cada cultura.

Culturas	NDVI	Exatidão Global	Parcelas Corretas	Percentagem de P. Suspeitas	Percentagem de P. Validadas
Milho	>0.4	>80%	627	0.64%	99.36%
Tomate	>0.3		389	6.17%	93.83%
Arroz	>0.3		303	3.63%	96.37%
Pastagens	>0.2		815	2.09%	97.91%

Ao observar a Tabela 3.10 verifica-se que, em termos percentuais, a cultura de tomate é a que apresenta maior valor de parcelas suspeitas.

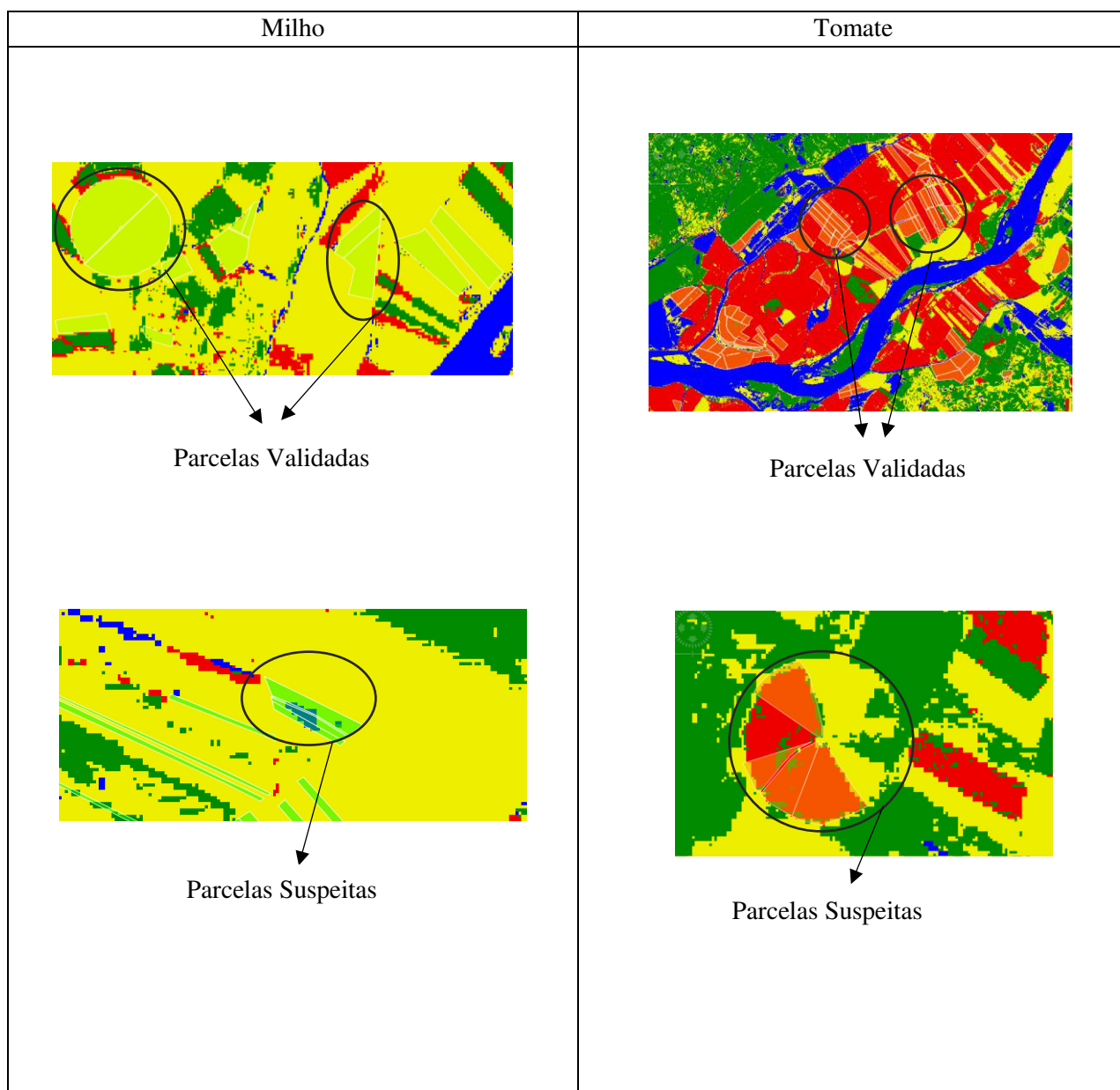


Figura 3.11: Exemplos de pormenores de zonas com parcelas de milho e tomate.

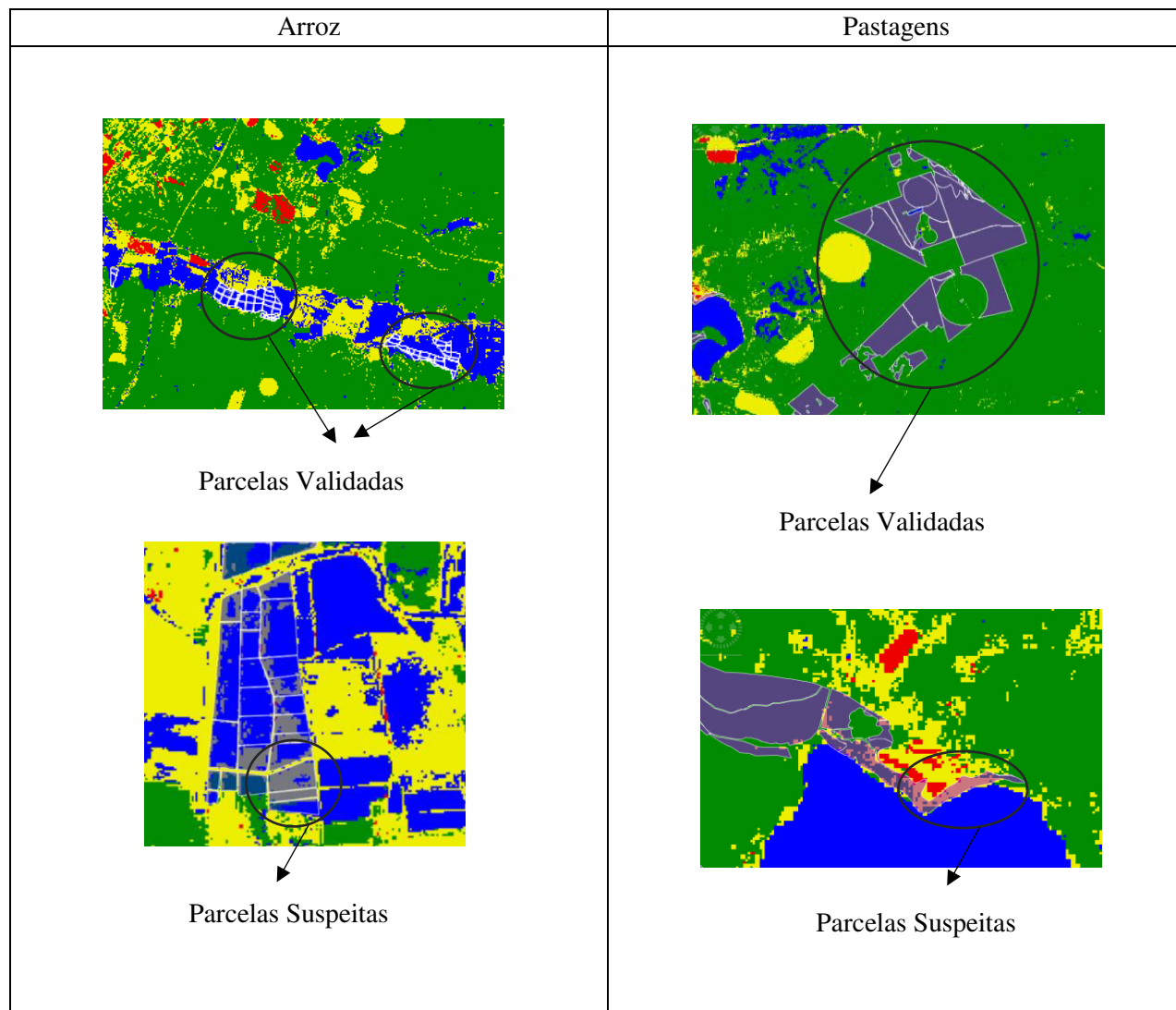


Figura 3.12: Exemplos de pormenores de zonas com parcelas de arroz e pastagens.

Nas Figuras 3.11 e 3.12 estão representados pormenores de exemplos de parcelas consideradas suspeitas e parcelas validadas para cada cultura. É possível ver para a cultura de arroz, nas parcelas consideradas suspeitas, zonas a cinzento que nos indicam, através da classificação, que essa zona possivelmente não terá arroz. O mesmo se verifica para a cultura de pastagens, onde na parcela suspeita se vê que possivelmente é uma zona de milho e não pastagens.

Para a cultura de milho o número de parcelas consideradas suspeitas foi bastante reduzido, sendo que as parcelas consideradas são aquelas que têm pequenas zonas onde a cor não é amarela, ou seja, pelo resultado da classificação não é considerada uma zona de milho. Relativamente à cultura de tomate, foi a que teve maior percentagem de parcelas suspeitas e também a cultura para a qual foi mais fácil, visualmente, verificar a existência de parcelas suspeitas (como a indicada na coluna direita da Figura 3.11) que nos mostra que apesar de classificada como tomate, é, sem dúvida, uma zona de milho.

4. Conclusões

Ao longo do presente projeto foram adquiridos conhecimentos sobre as metodologias utilizadas no IFAP e principalmente sobre a metodologia utilizada na parte do parcelário para identificação e controlo agrícola. O objetivo do projeto passou por avaliar o ficheiro fornecido pelo IFAP que continha variados tipos de culturas, das quais foram estudadas quatro culturas do controlo físico do ano de 2016, sendo elas milho, arroz, tomate e pastagens.

Foram recolhidas imagens de satélite que abrangiam o ciclo fenológico das culturas de forma a ser possível estudar cada uma através do comportamento visível pela resposta espectral das imagens. A partir de amostras para cada cultura obteve-se, com recurso ao NDVI, valores de média e desvio padrão utilizados ao longo do estudo. Também foram calculados outros oito índices de vegetação, sendo que a grande diferença encontrada, comparativamente ao NDVI, esteve na estabilidade apresentada no índice CIRedEdge no início e no fim da série temporal estudada. As culturas que melhor se distinguem através do valor NDVI foram o milho e o tomate, apesar de ser possível diferenciar bastante bem as pastagens das outras três culturas em estudo, devido ao seu comportamento distinto.

Foram identificadas as declarações que efetivamente têm veracidade na sua declaração, obtendo assim o número de declarações verdadeiras para cada cultura. De salientar que, após a análise das séries temporais do NDVI para cada cultura, foi possível identificar as parcelas que apresentavam um comportamento diferente do esperado, tendo estas sido utilizadas, em seguida, para se efetuar a classificação das imagens com recurso a áreas de treino e validação corretas.

Posteriormente foi avaliada a média das médias do NDVI para cada cultura, a partir de onde foi possível encontrar as melhores datas que evidenciavam diferenças entre as culturas. Escolhidas estas datas passou-se para a classificação de imagens utilizando três classificadores e três conjuntos de classes. Concluiu-se que o melhor classificador é o classificador de Redes Neurais, tanto para a imagem das 9 bandas como para a imagem dos índices. Os testes iniciais foram todos feitos com o classificador de Máxima Verosimilhança por se apresentar mais rápido, sendo que apenas os conjuntos que apresentaram melhores resultados foram então testados com o classificador de Redes Neurais.

A maior dificuldade encontrada neste projeto consistiu na escolha do valor limite para excluir as parcelas que não se encontravam cultivadas, uma vez que ao escolher um valor demasiado alto poderiam ser eliminadas parcelas corretamente controladas, em oposição ao facto de ao escolher um valor baixo poder incluir parcelas erradamente controladas. A não diferenciação das culturas de milho precoce e milho tardio e também de tomate precoce e tomate tardio (visível a diferença do comportamento entre estas culturas no gráfico da Figura 3.10) pode levar a que ocorra um erro na validação, assumindo que existe uma parcela de milho que não o é na realidade. O contrário também se pode verificar, existindo uma parcela que contém milho, mas é considerada como parcela suspeita.

A validação das parcelas mostrou resultados bastante positivos, apesar de existirem parcelas controladas erradamente. Também a classificação de imagens permitiu encontrar um método de controlo agrícola através de Detecção Remota muito mais eficaz do que o controlo feito atualmente pelo IFAP. Assim, o objetivo proposto foi conseguido e os resultados satisfatórios.

Trabalho futuro:

Quanto a perspetivas futuras com vista ao melhoramento destas metodologias, salienta-se a automatização do processo de deteção de falsas declarações e a caracterização de outros tipos de culturas de forma a ter uma maior diversidade fenológica, e avaliar em que medida tal beneficiaria a sua diferenciação. Para além disso, a utilização de séries temporais com intervalos de tempo mais curtos, com recurso a imagens dos 2 satélites da missão Sentinel-2, pode levar a que exista uma melhor discriminação das diferentes culturas devido ao facto de permitirem a criação de séries temporais mais completas.

Referências

- Baio, M. e Matos, J. (2007): “Detecção Remota”. Texto não publicado, Instituto Superior Técnico, Lisboa.
- CCRS (2017): “Fundamentals of Remote Sensing”: http://www.nrcan.gc.ca/sites/www.nrcan.gc.ca/files/earthsciences/pdf/resource/tutor/fundam/pdf/fundamentals_e.pdf, acedido em Maio 2017.
- Catalão, J. (2012): “Apontamentos de Princípios e Aplicações da Detecção Remota.” Texto não publicado, Faculdade de Ciências da Universidade de Lisboa, Lisboa.
- Clevers, J.G.P.W. e Gitelson A.A. (2013): “Remote estimation of crops and grass chlorophyll and nitrogen content using red-edge bands on Sentinel-2 and -3.” *International Journal of Applied Earth Observation and Geoinformation*, Vol. 23, pp. 344-351.
- DGT (2017): “Sentinel-2”: http://www.dgterritorio.pt/cartografia_e_geodesia/cartografia/detecao_remota/copernicus/componente_espaco/breve_descricao_dos_satelites_sentinel/, acedido em Maio 2017.
- ESA (2017): <https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/resolutions/spatial>, acedido a Julho 2017.
- ESA (2017): <https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/resolutions/spatial>, acedido a Julho 2017.
- Fonseca, A. D. e Fernandes J.C. (2004): “Detecção Remota”, LIDEL – Edições Técnicas, Lisboa, Portugal.
- Hussain, M., D. Chen, A. Cheng, H. Wei e D. Stanley (2013) “Change detection from remotely sensed images: From pixel-based to object-based approaches.” *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 80, pp. 91-106.
- IFAP (2017): http://www.ifap.min-agricultura.pt/portal/page/portal/ifap_publico/GC_oifap#.WVNhN-vyvIU, acedido em Maio 2017.
- Karjalainen, M.; Kaartinen, H.; Hyypä, J. (2008): “Agricultural Monitoring Using Envisat Alternating Polarization SAR images”, *Photogrammetric Engineering & Remote Sensing*, Vol. 74, págs. 117-126.
- Khan, M.R., C.A.J.M. de Bie, H. van Keulen, E.M.A. Smalling e R. Real (2010): “Disaggregating and mapping crop statistics using hypertemporal remote sensing.” *International Journal of Applied Earth Observation and Geoinformation*, Vol. 12, Nº1, pp. 36-46.
- Lu, D., P. Mausel, E. Brondízio e E. Moran (2004): “Change detection techniques.” *International Journal of Remote Sensing*, Vol. 25, Nº12, pp. 2365-2407.
- McNairn, H.; Champagne, C.; Shangmstrom, D.; Reichert, G. (2009): “Integration of optical and Synthetic Aperture Radar (SAR) imagery for delivering operational annual crop inventories”, *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 64, No 5, págs. 434-449.

Peña-Barragón J.M., Ngugi M.K., Plant R.E. e Six J. (2011): “Object-based crop identification using multiple vegetation indices, textural features and crop phenology.” *International Journal of Applied Earth Observation and Geoinformation*, Vol. 115, Nº6, pp. 1301-1316.

Richards, J.A. e X. Jia (2006) *Remote Sensing Digital Image Analysis*, 4.^a edição, Springer-Verlag Berlin Heidelberg, Alemanha, pp. 193-263.

Sousa, A. e Silva, J. (2011): “Fundamentos Teóricos da Deteção Remota”. Texto não publicado, Universidade de Évora, Évora.

Zikkmann, E. e H. Weichelt (2014): “Crop identification by means of seasonal statics of RapidEye time series.”, *Third International Conference on Agro-Geoinformatics*, Center for spatial information science and system (CSISS), Geoge Mason University and Institute of Agricultural Resources and Regional Planning, Beijing, China, 11-14, Agosto 2014, pp. 26-31.

Anexos

Anexo I – Resultados do Classificador de Máxima Verosimilhança

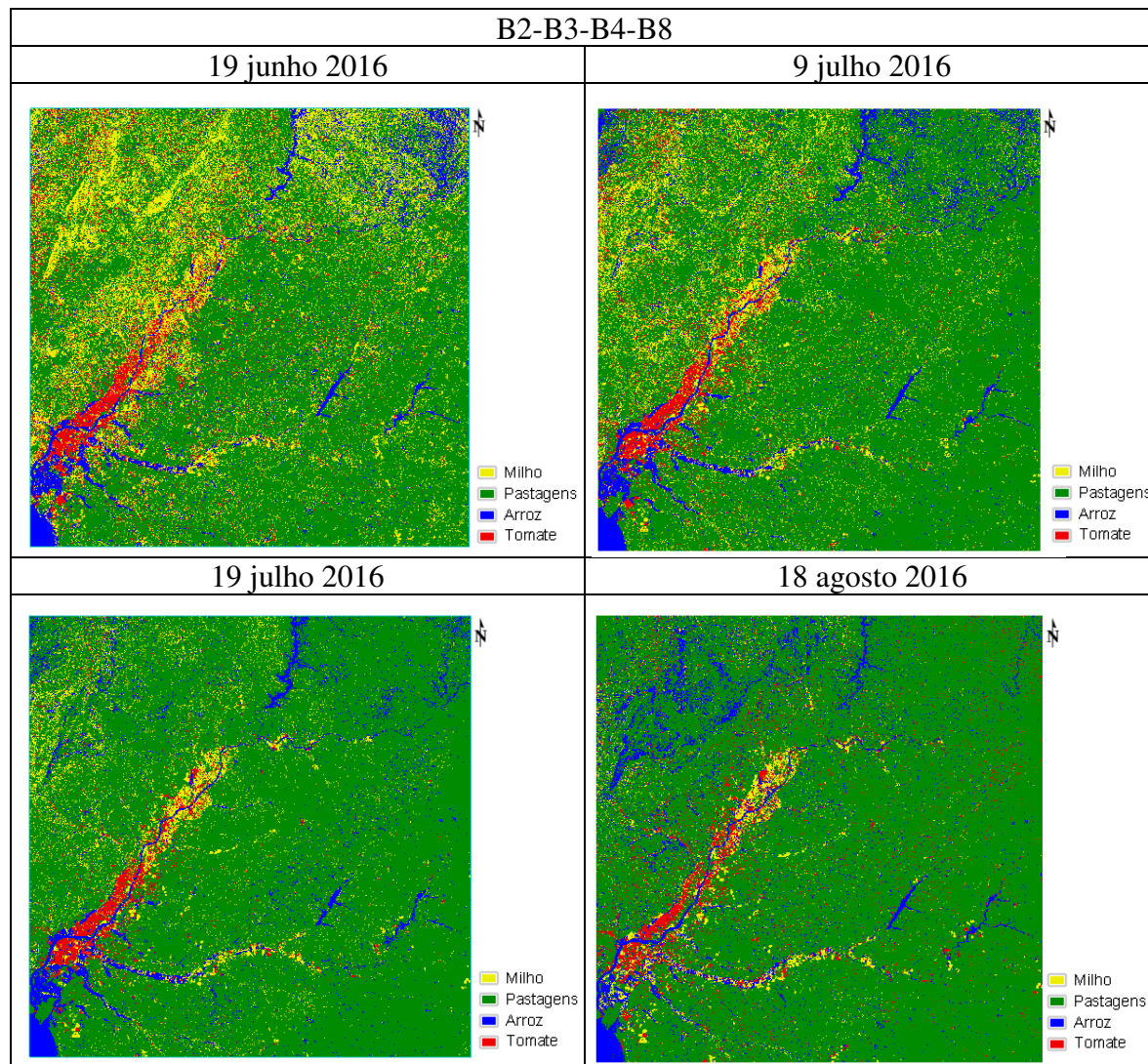


Figura Anexo 1: Imagens com 4 bandas para as 4 datas escolhidas para a validação com 4 classes utilizando o Classificador de Máxima Verosimilhança.

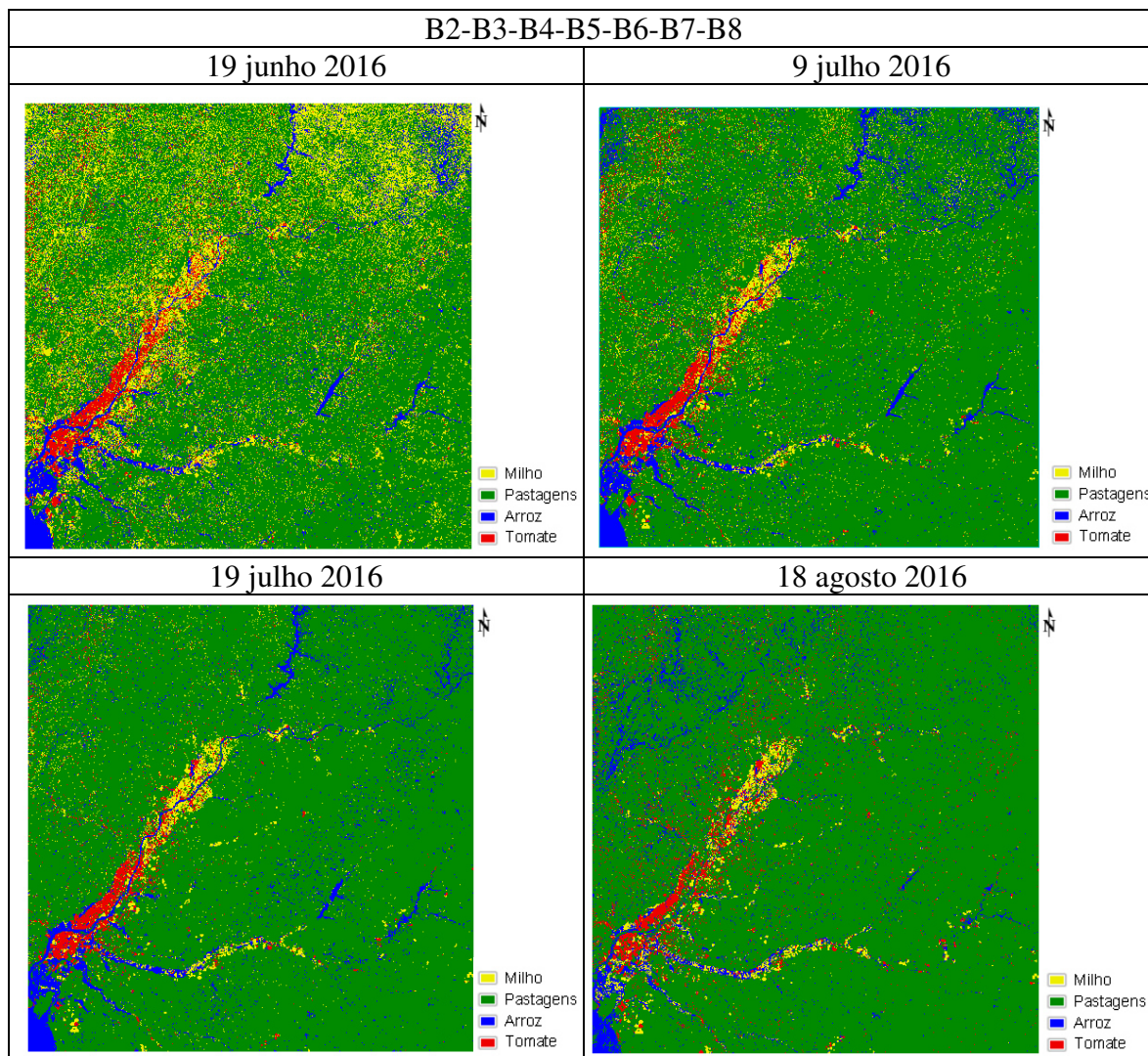


Figura Anexo 2: Imagens com 7 bandas para as 4 datas escolhidas para a validação com 4 classes utilizando o Classificador de Máxima Verossimilhança.

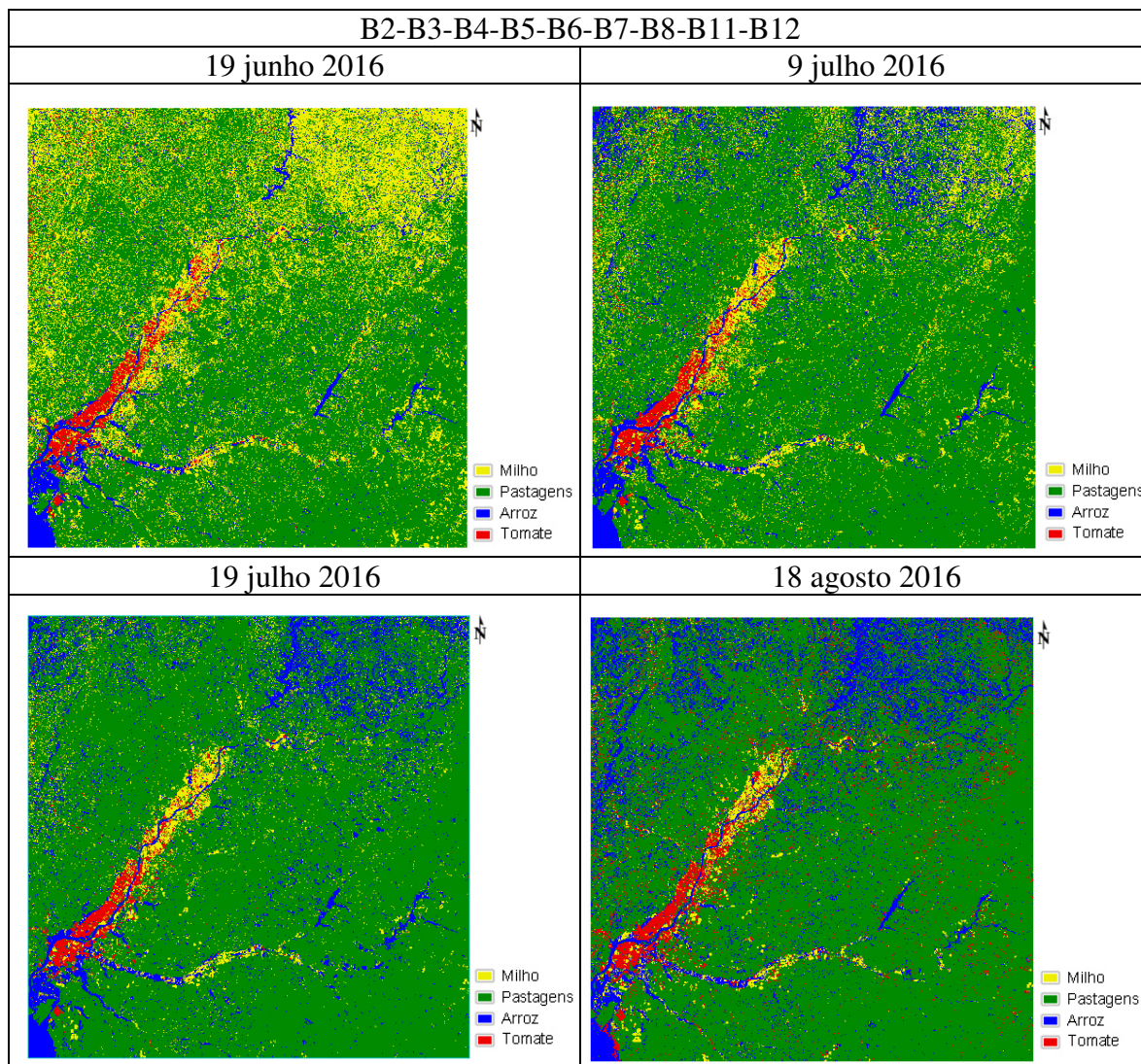


Figura Anexo 3: Imagens com 9 bandas para as 4 datas escolhidas para a validação com 4 classes utilizando o Classificador de Máxima Verossimilhança.

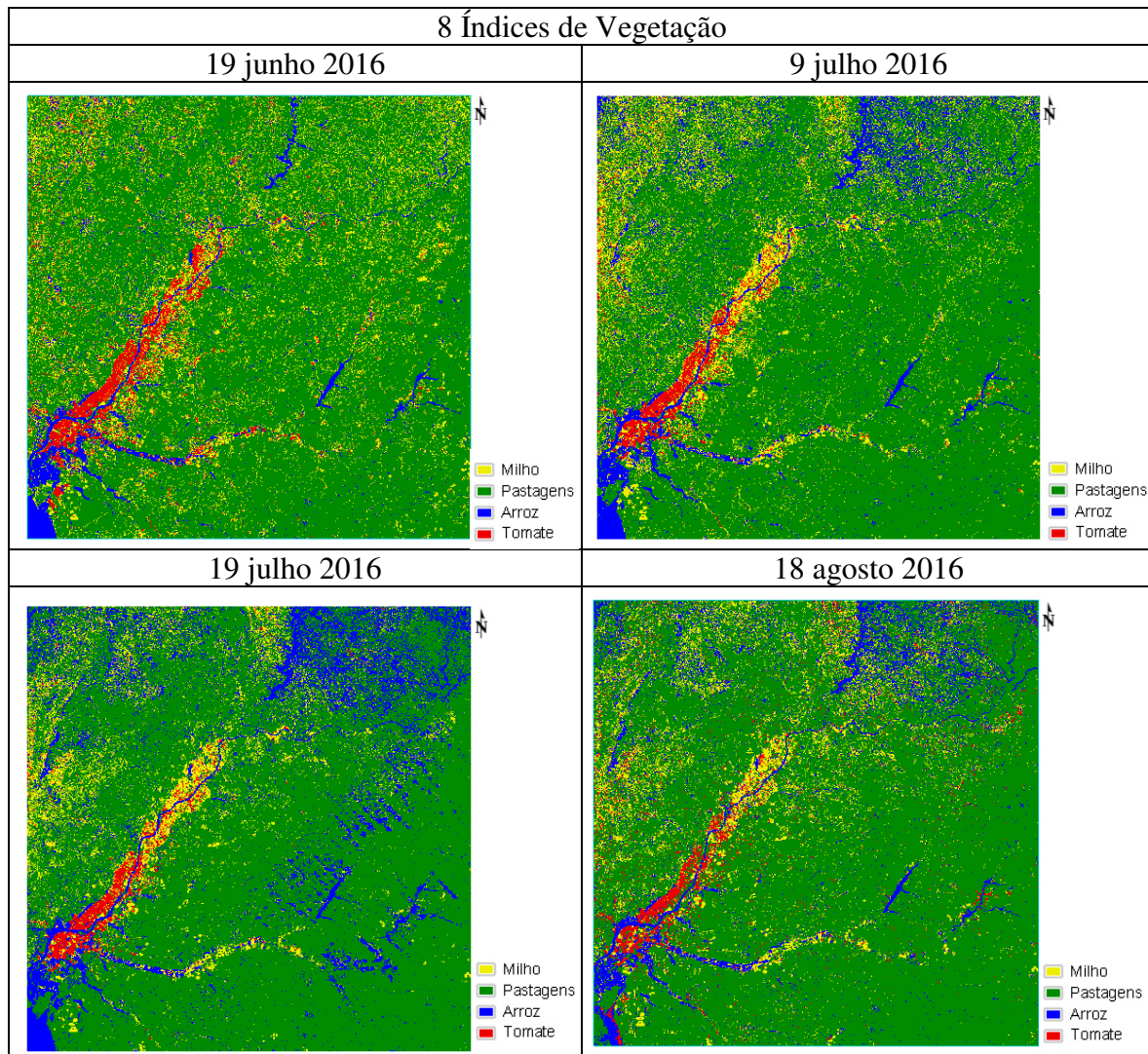


Figura Anexo 4: Imagens com 8 índices para as 4 datas escolhidas para a validação com 4 classes utilizando o Classificador de Máxima Verossimilhança.

Anexo II – Resultados do Classificador de Redes Neurais

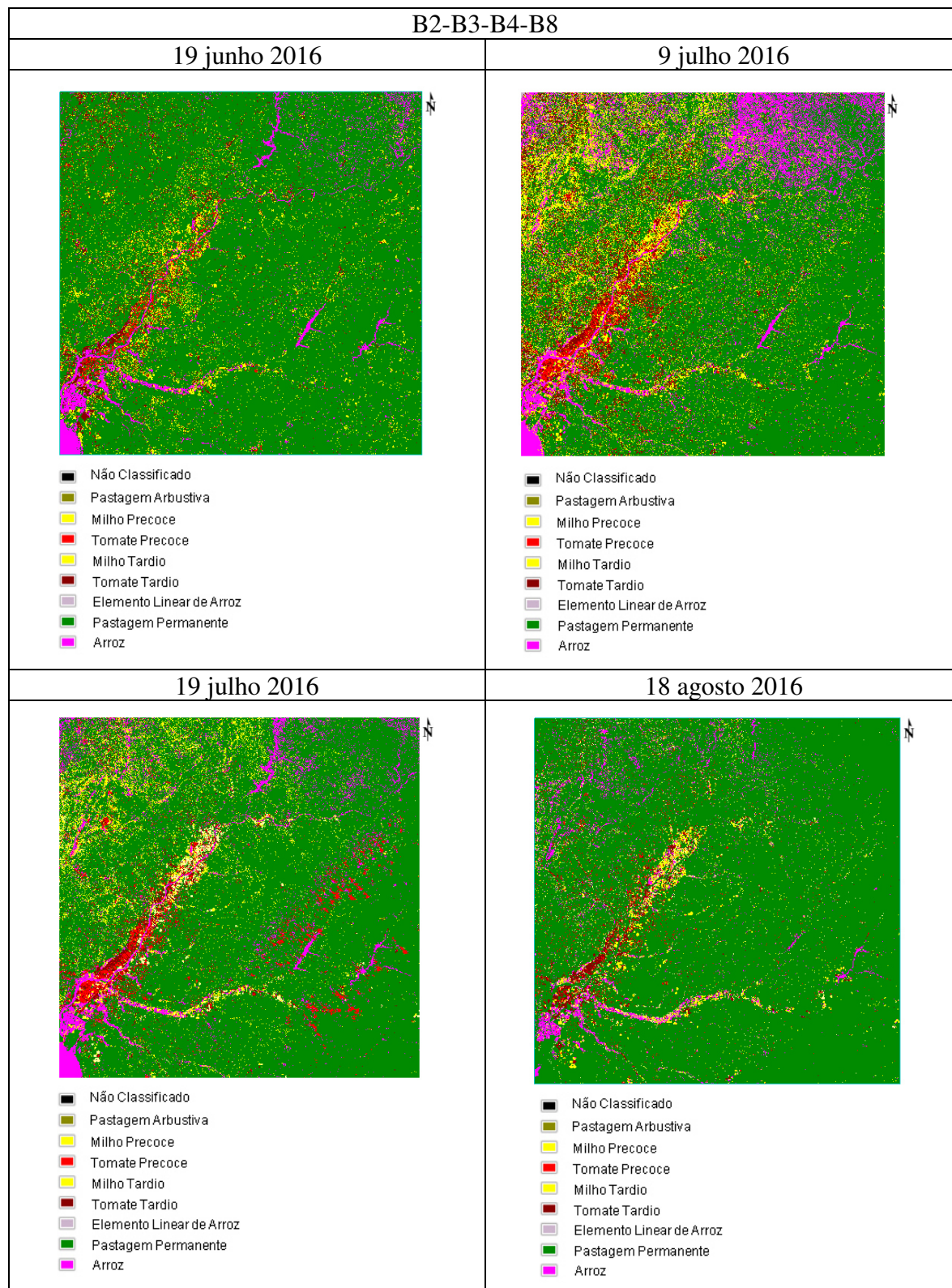


Figura Anexo 5: Imagens com as 4 bandas para as 4 datas escolhidas para a validação com 8 classes utilizando o Classificador de Redes Neurais.

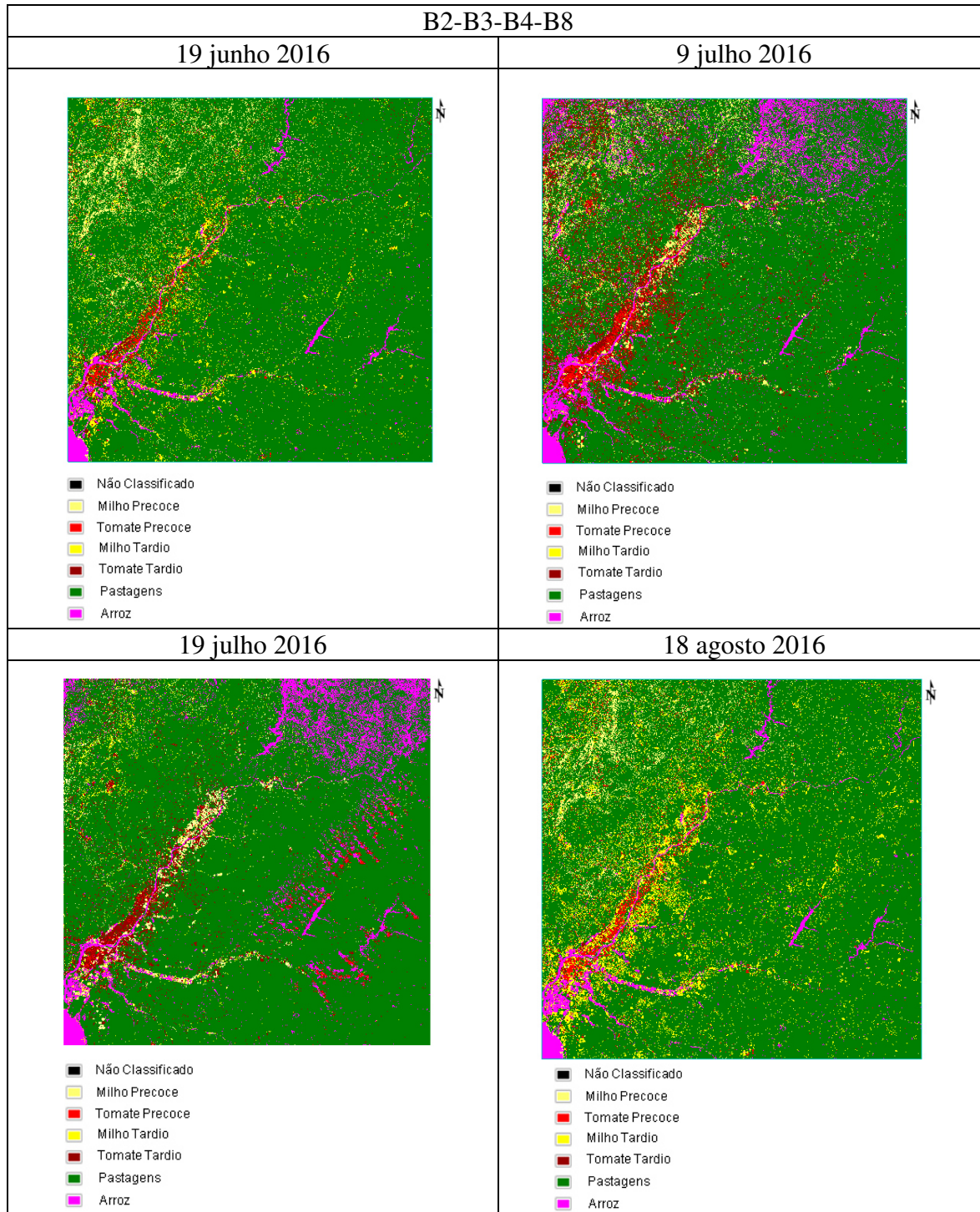


Figura Anexo 6: Imagens com as 4 bandas para as 4 datas escolhidas para a validação com 8 classes utilizando o Classificador de Redes Neurais.

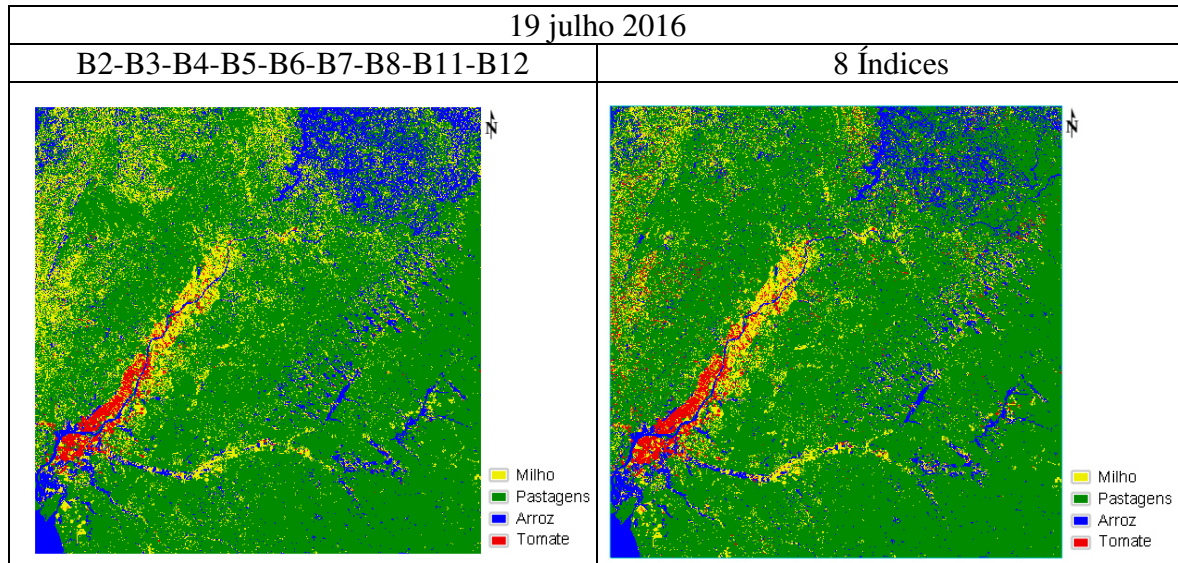


Figura Anexo 7: Imagens com as 9 bandas e com os 8 índices para a melhor data da validação com 4 classes utilizando o Classificador de Redes Neurais.

Anexo III – Índices de vegetação

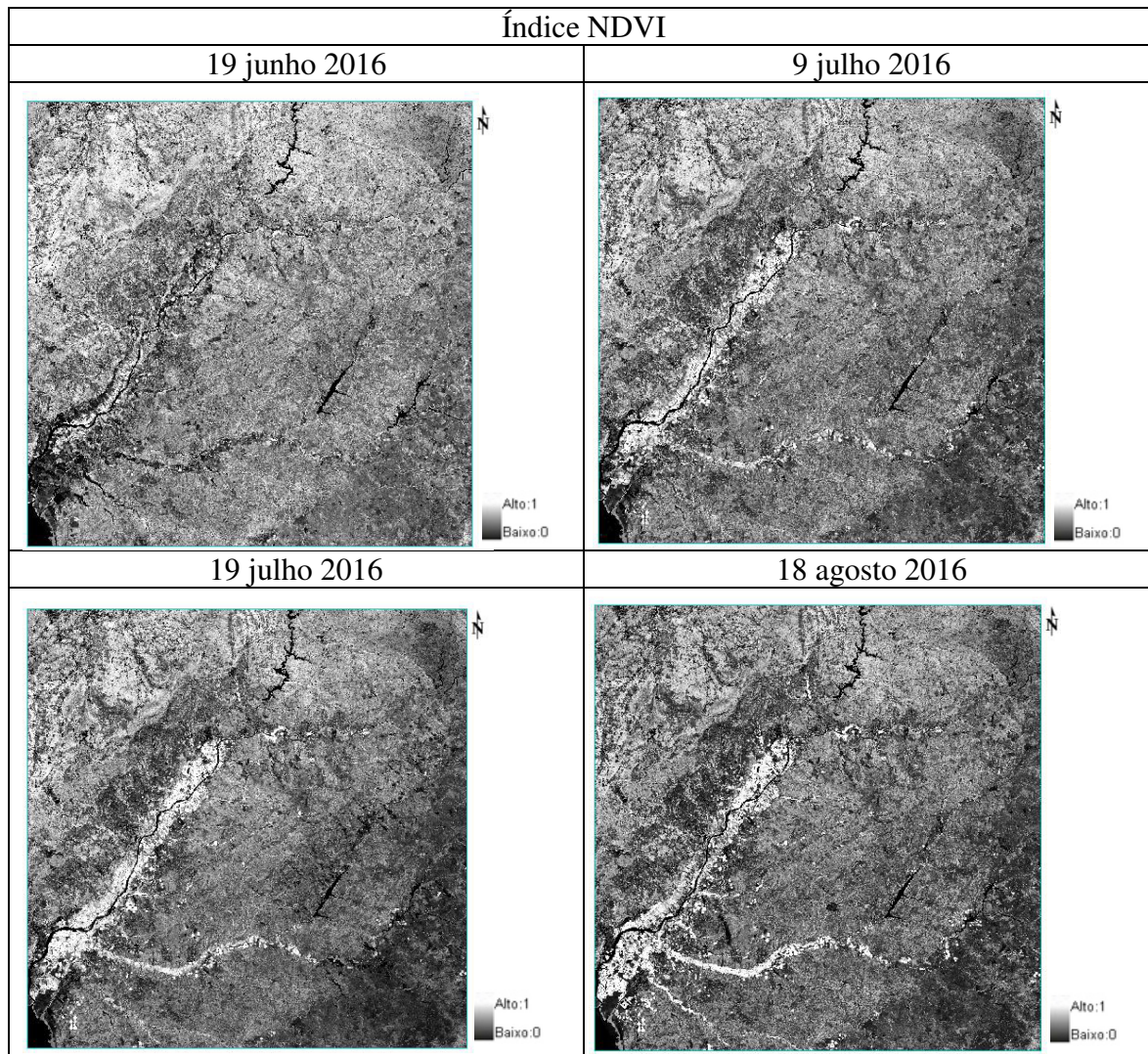


Figura Anexo 8: Imagens para o índice NDVI.

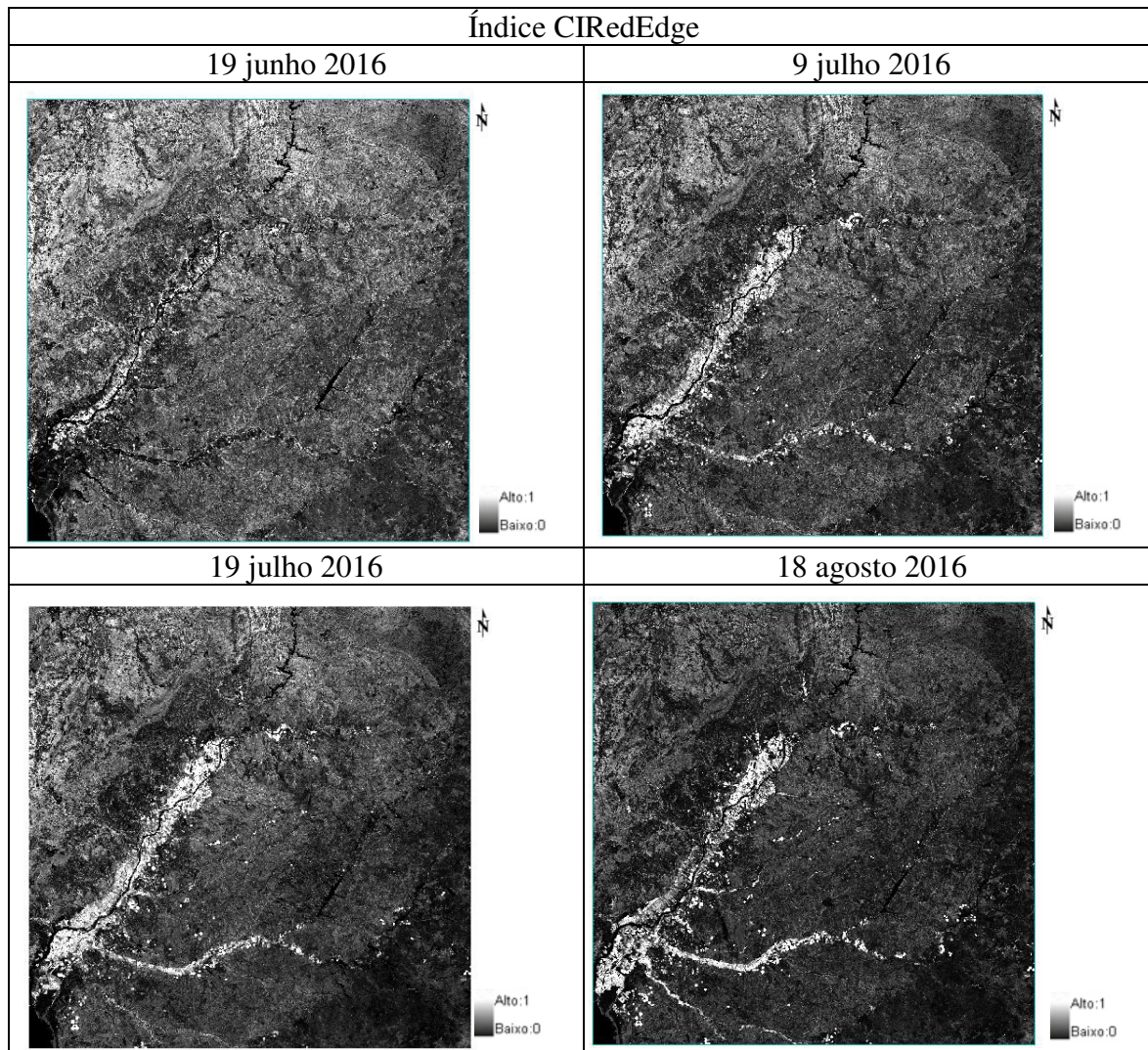


Figura Anexo 9: Imagens para o índice CIRedEdge.

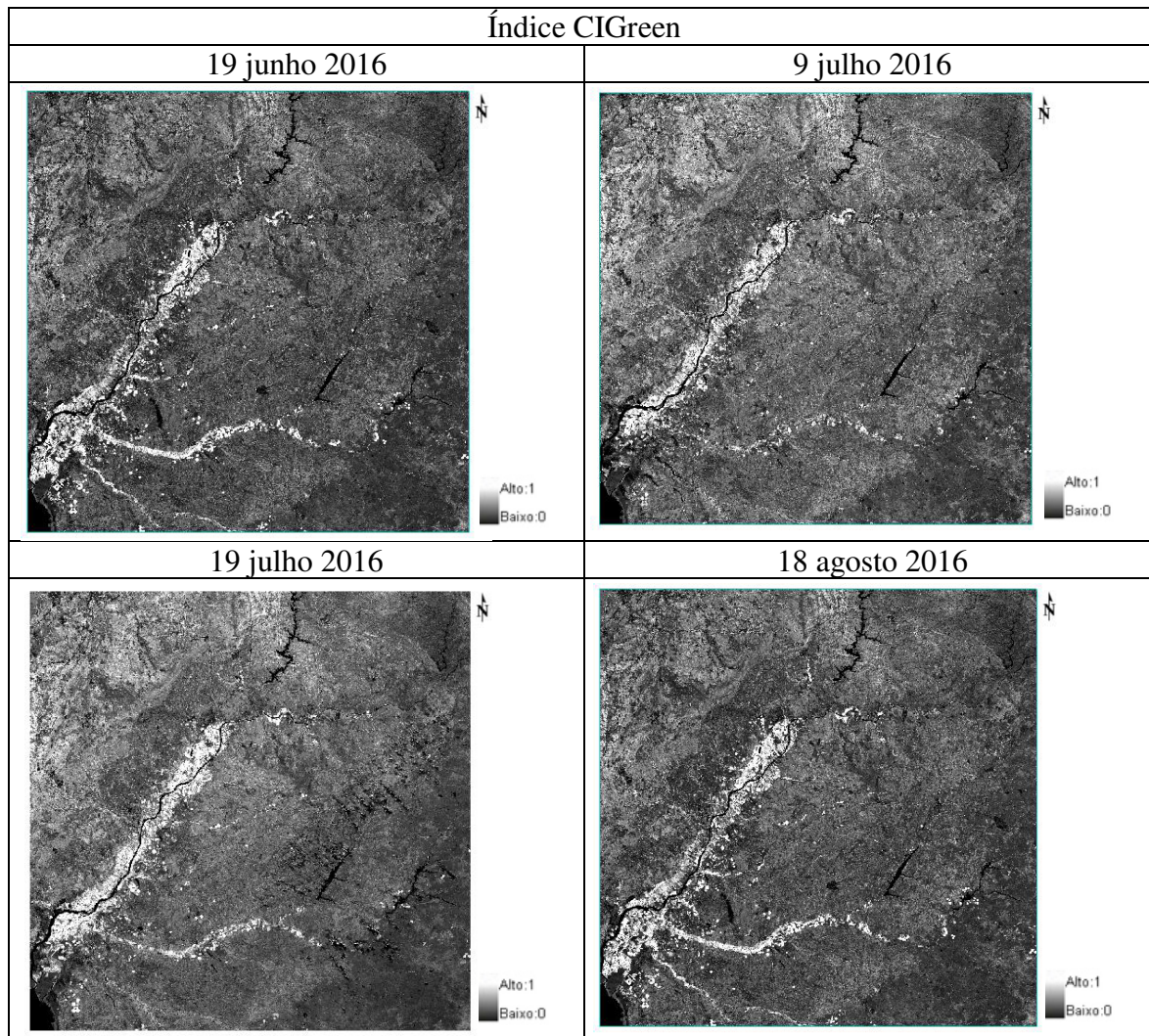


Figura Anexo 10: Imagens para o índice CIGreen.

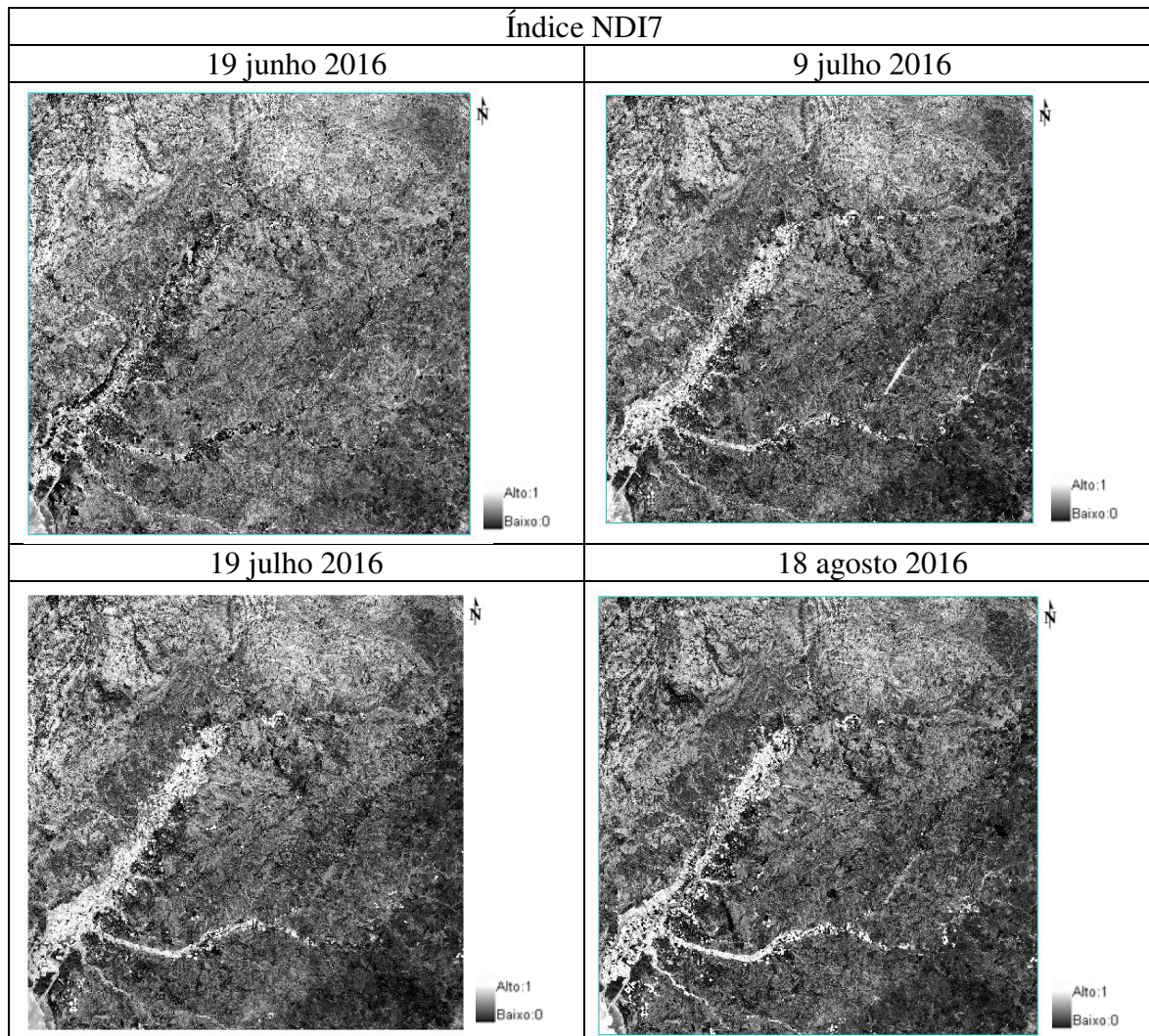


Figura Anexo 11: Imagens para o índice NDI7.

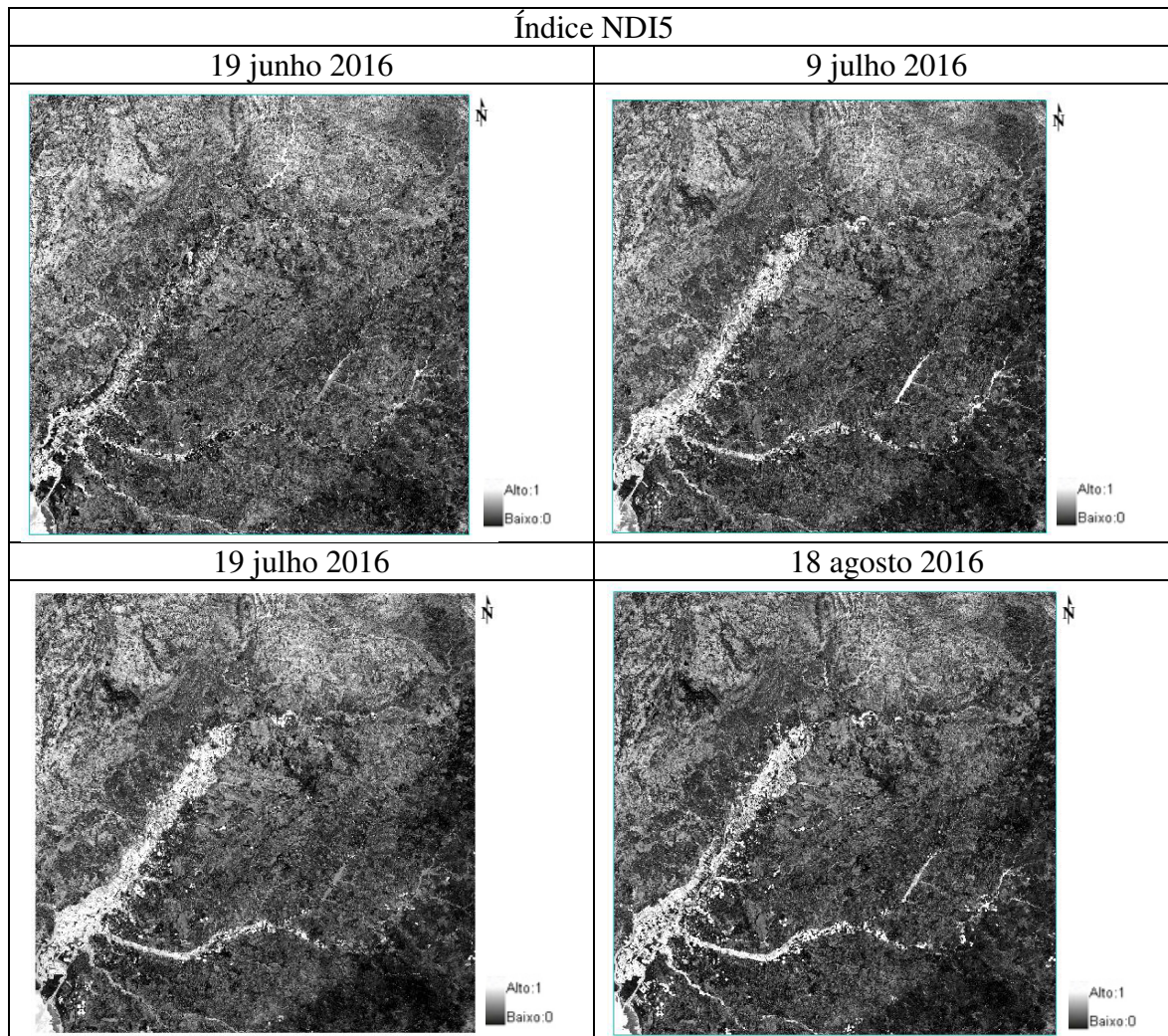


Figura Anexo 12: Imagens para o índice NDI5.

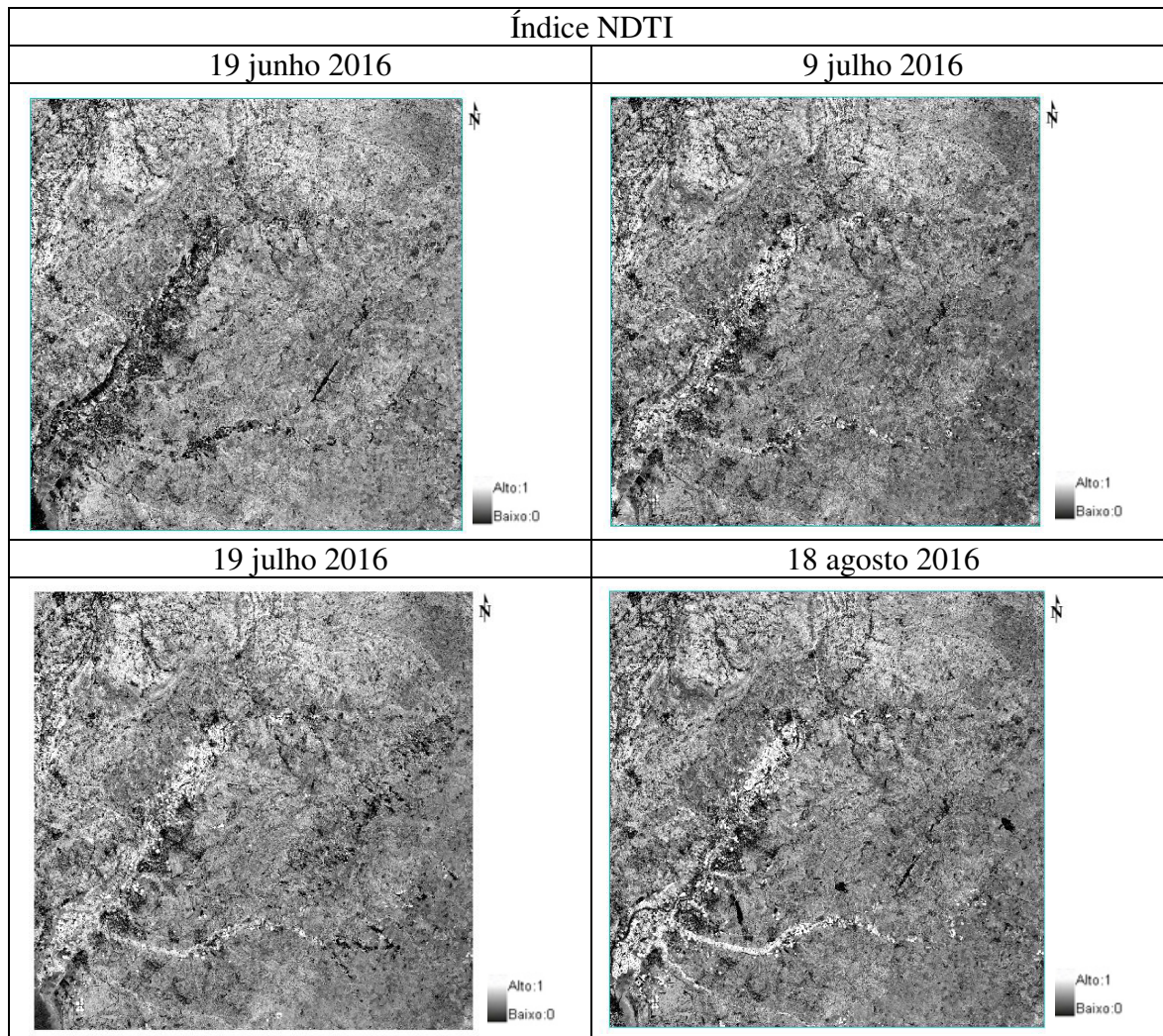


Figura Anexo 13: Imagens para o índice NDTI.

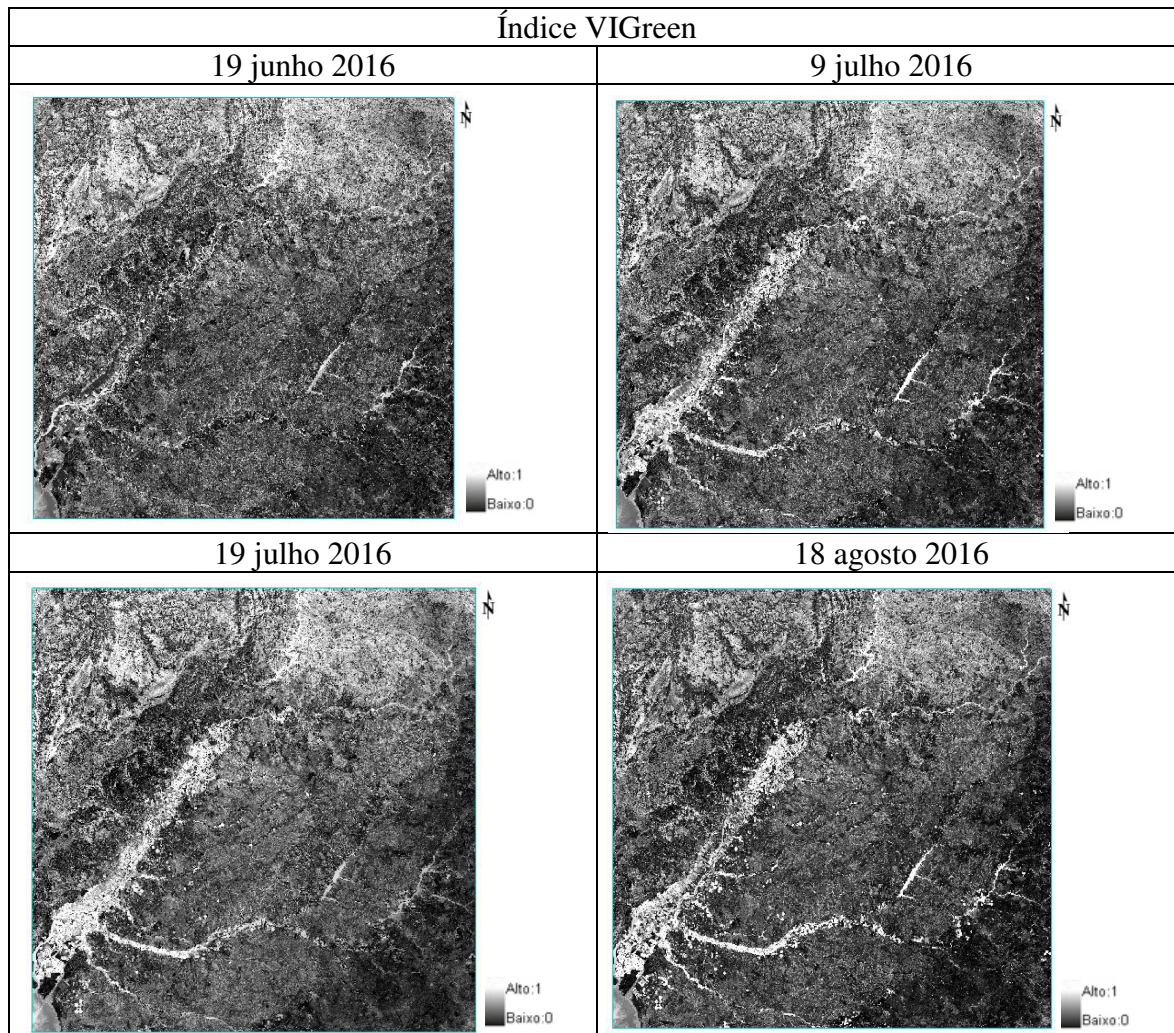


Figura Anexo 14: Imagens para o índice VIGreen.

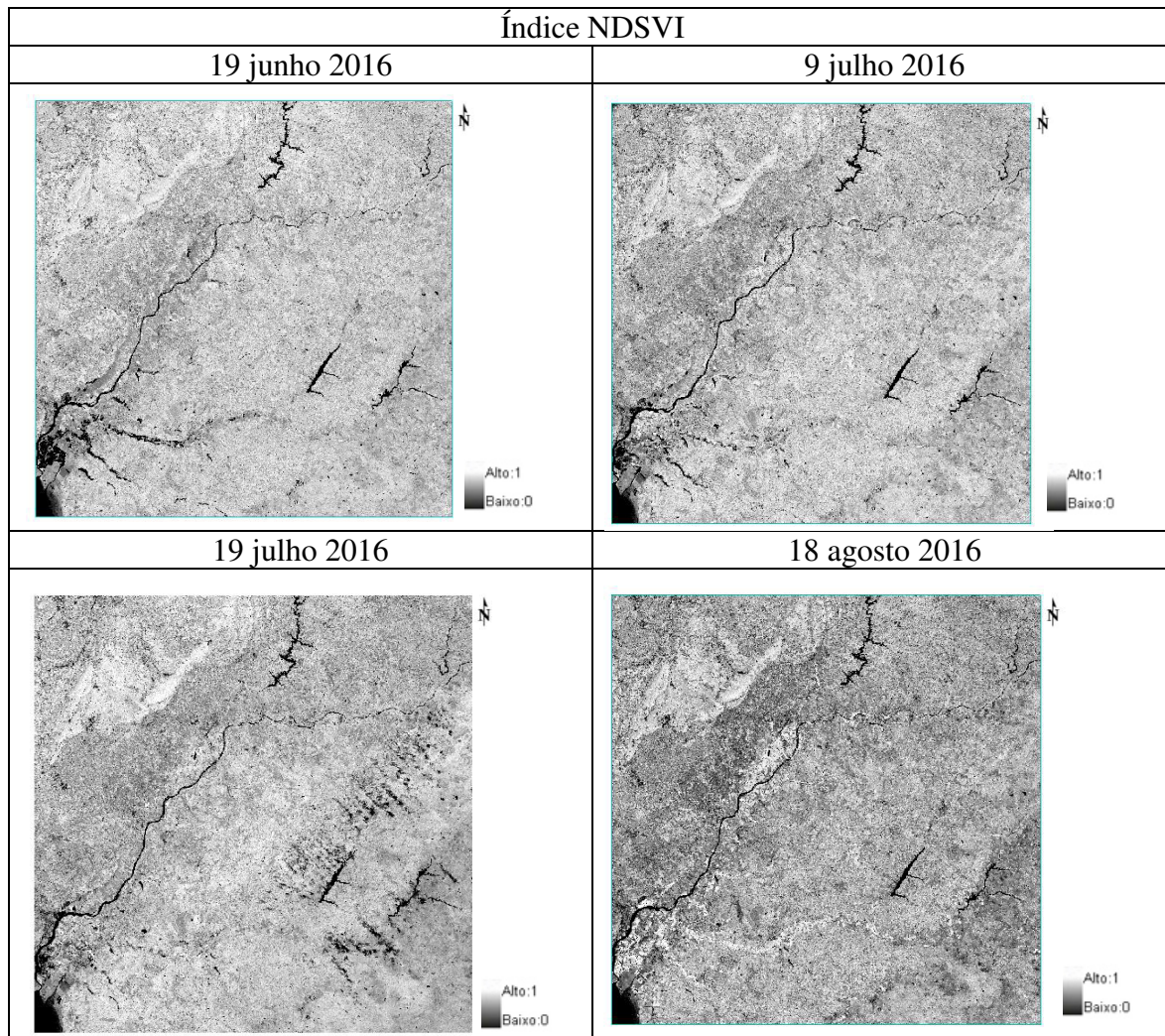


Figura Anexo 15: Imagens para o índice NDSVI.

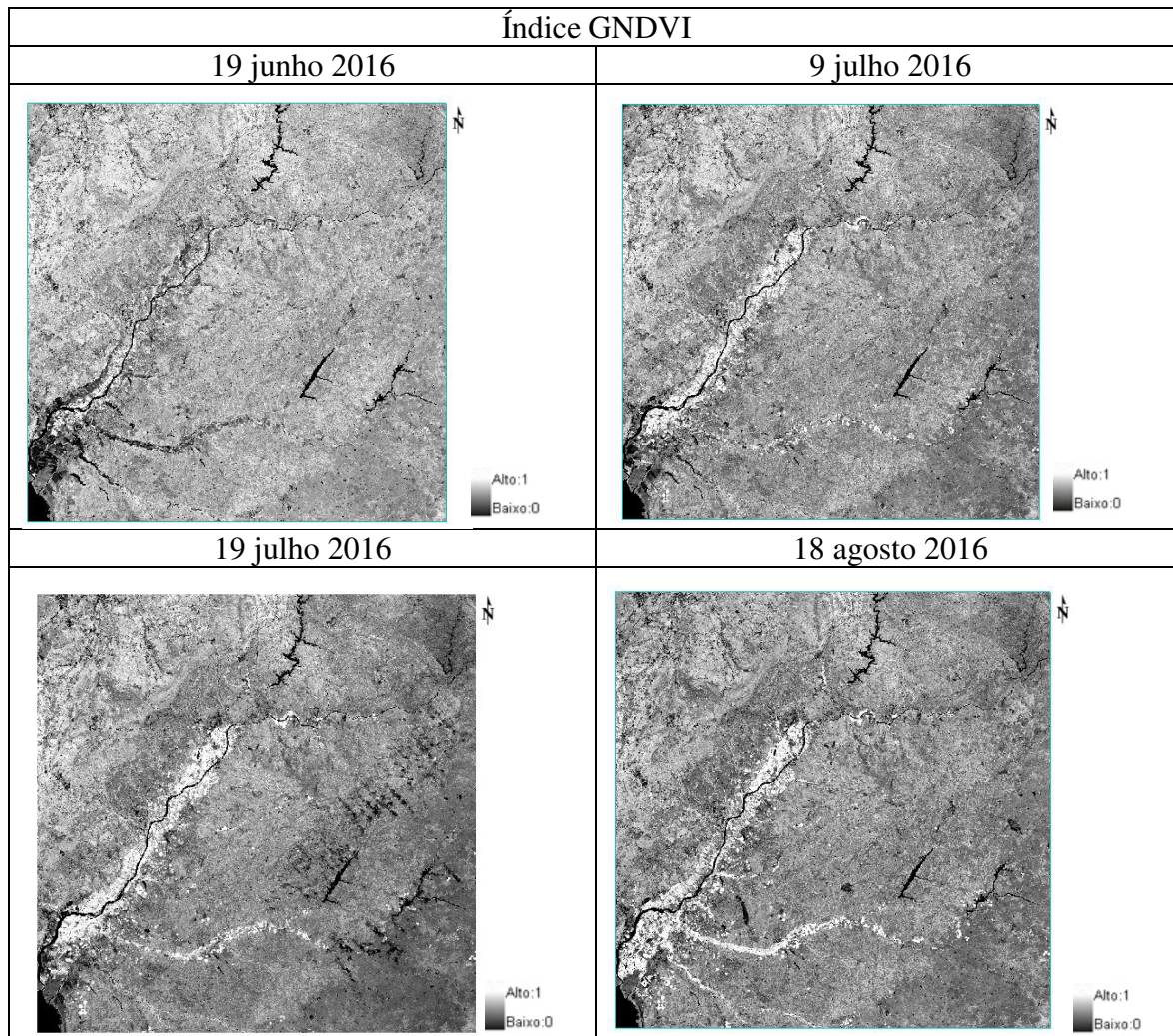


Figura Anexo 16: Imagens para o índice GNDVI.

Anexo IV – Parcelas das culturas

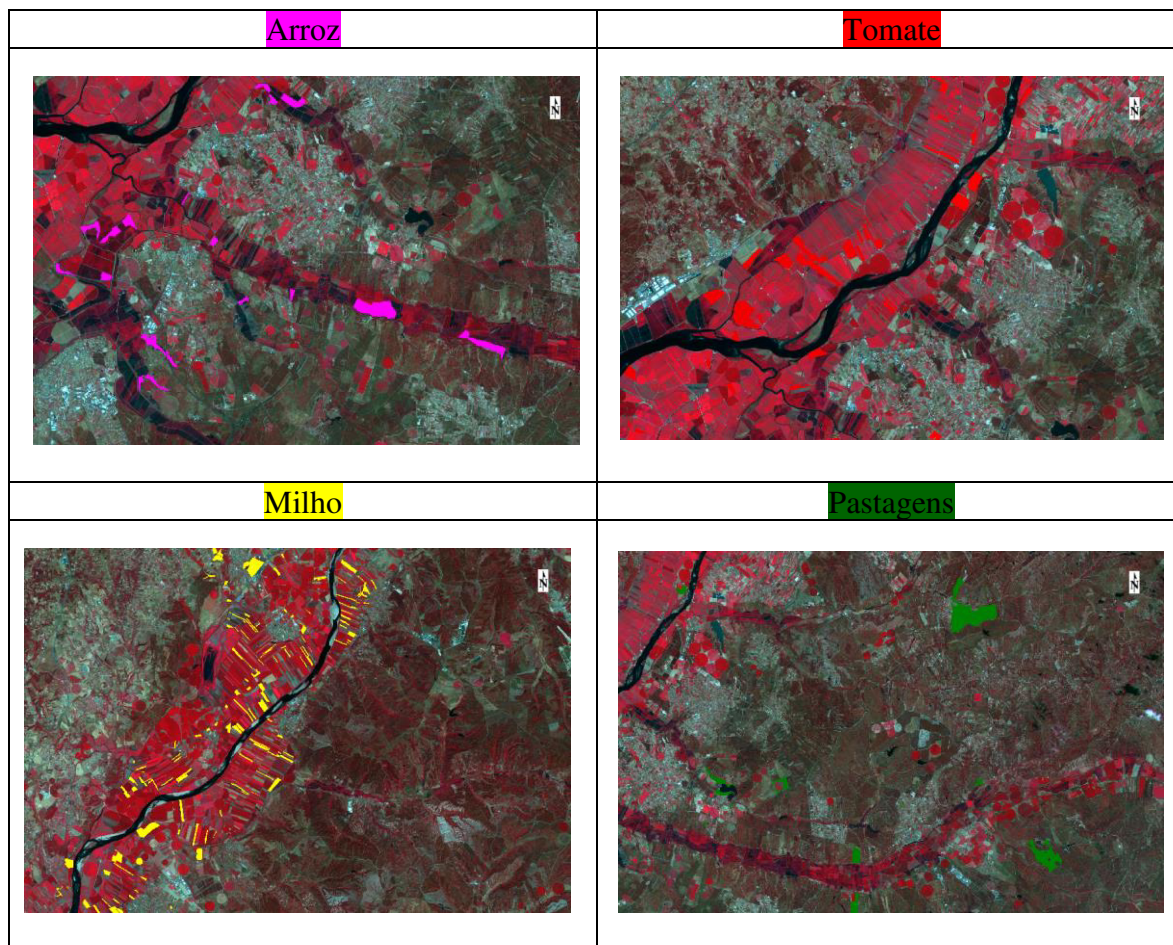


Figura Anexo 17: Zoom de zonas com maior incidência de cada cultura.

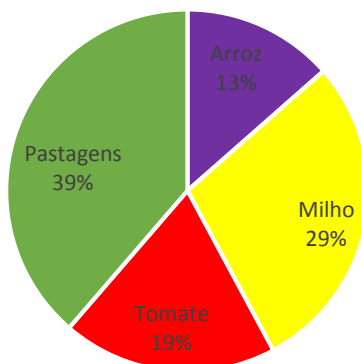


Figura Anexo 18: Percentagem de cada cultura existente no estudo.