

Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Identificação automática de zonas de floresta e de alteração de floresta utilizando plataformas disponíveis e dados do *Sentinel 2*

Ana Isabel dos Santos Ferreira

Dissertação de Mestrado na área científica de Engenharia de Informação Geoespacial orientada pela Professora Doutora Cidália Maria Parreira da Costa Fonte e coorientada pelo Professor Doutor Diogo André Vicente Amorim Duarte e apresentada ao Departamento de Matemática da Faculdade de Ciências e Tecnologia da Universidade de Coimbra.

Setembro de 2020



UNIVERSIDADE D
COIMBRA

Índice

Índice.....	i
Agradecimentos.....	iii
Resumo.....	iv
Abstract.....	v
Acrónimos, Siglas e Símbolos	vi
Índice de Figuras, Tabelas e Fluxogramas.....	ix
1. Introdução.....	1
1.1 Enquadramento e motivação	2
1.2 Objetivos.....	3
1.3 Organização da dissertação	3
2. Conceitos de Detecção Remota.....	5
2.1. A missão <i>Sentinel 2</i>	5
2.2. Classificação de imagens multiespectrais	8
2.3. Índices radiométricos.....	9
2.4. Validação de mapas de ocupação do solo	12
3. Portais de monitorização Florestal	15
3.1. <i>Sentinel Hub</i>	15
3.2. <i>Terramonitor</i>	16
3.3. <i>Giovanni</i>	17
3.4. <i>Land Cover</i>	19
3.5. <i>Global Forest</i>	21
3.6. <i>F-Tep</i>	23
3.7. Comparação das plataformas.....	25
4. Portal F-Tep.....	27
4.1. Produtos disponíveis.....	27
4.1.1. Para imagens do Sentinel-2	27

4.1.1.1.	Serviço LandCoverS2.....	28
4.1.1.2.	Serviço VegetationIndices.....	29
4.1.1.3.	Serviço VegetationIndices OLD.....	30
4.1.1.4.	Serviço ForestChangeS2.....	30
4.1.1.5.	Serviço CombS2granules & CombS2granulesZ.....	31
4.1.2.	Para imagens do Sentinel-1	33
4.1.2.1.	Serviço LandCoverS1.....	33
4.1.2.2.	Serviço S1Biomass.....	34
4.1.2.3.	Serviço S1stack.....	35
4.1.3.	Para imagens de qualquer satélite.....	35
4.1.3.1.	Serviço LandCoverGeotiff.....	36
5.	Identificação automática das zonas de floresta e de alteração de floresta	37
5.1.	Caso de Estudo	37
5.2.	Dados de referência	39
5.3.	Dados de treino	40
5.4.	Mapas de ocupação do solo	44
5.5.	Máscara NBR	46
5.6.	Máscara da floresta.....	53
5.6.1.	Validação da máscara da floresta	55
5.7.	Alterações na classe floresta.....	58
5.7.1.	Validação da máscara de alterações da floresta.....	60
6.	Discussão & Conclusões.....	63
6.1.	Identificação automática de florestas.....	64
6.2.	Identificação automática de alterações nas florestas	66
6.3.	Limitações & Trabalho futuro	66
	Referências	68

Agradecimentos

Várias são as pessoas responsáveis pela conclusão deste projeto e conseqüentemente pela conclusão do meu Mestrado em Engenharia de Informação Geoespacial. Em primeiro lugar, e como não poderia deixar de ser, quero agradecer aos meus orientadores Professora Doutora Cidália Fonte e Professor Doutor Diogo Duarte por toda a ajuda prestada, por todo o conhecimento transmitido, por todo o empenho e dedicação neste projeto e por todo o seu profissionalismo demonstrado como orientadores de uma Tese de Mestrado. Um obrigada por tudo!

Em seguida gostaria de agradecer ao colaborador do meu projeto, o instituto de Investigação da Floresta e do Papel (RAIZ), por me oferecer as condições necessárias para a realização deste projeto. Como não poderia deixar de ser, quero agradecer em especial à Doutora Margarida Silva e ao Mestre André Duarte por estarem sempre disponíveis e disponibilizarem a sua ajuda sempre que necessário.

Por fim, aos mais importantes e sempre presentes nesta minha caminhada académica pela cidade dos amores: a família e os amigos. Em relação à família, quero destacar os meus pais e o meu irmão pelo apoio e ajuda incondicional. Quanto aos amigos quero agradecer especialmente: ao Edgar pela sua compreensão e cumplicidade, ao Ismael, ao Tomás, à Cata, à Catarina, à Patrícia, à Liliana, à Jéssica, à Diana e a todos que de certa forma me ajudaram nesta jornada, sem eles nada disto seria possível.

Resumo

Este trabalho foi realizado com enquadramento do projeto *Exploring the Potential of the Sentinel missions Satellite Imagery* (EPSSI) e feito em estreita colaboração com o RAIZ, baseou-se essencialmente na criação de mapas de ocupação do solo, com particular enfoque na floresta. Foram testadas algumas ferramentas disponíveis no portal *Forestry Thematic Exploitation Platform* (F-Tep) para a identificação automática das zonas de floresta e de alteração de floresta. Todos estes objetivos foram tomados tendo em conta a importância do conhecimento da floresta no nosso território, que tipo de florestas estão disponíveis, em que quantidade e de que modo se pode gerir os seus materiais de forma sustentável.

Inicialmente foi feito um estudo dos produtos disponíveis para fins florestais em outras entidades, para além do RAIZ, tais como: *Sentinel Hub*, *Terramonitor*, *F-Tep*, *Giovanni*, *Land Cover* e *Global Forest*. Só podemos saber se um produto nos será útil quando conhecemos as suas especificidades, para isso, foram estudadas as características dos produtos disponíveis em cada portal: o período de atualização, o custo, a resolução espacial, a nomenclatura, a fonte de dados e como são obtidos os produtos. Feita a análise dos produtos existentes em todos os portais foram selecionados os que seriam úteis para uma instituição como o RAIZ. Notou-se que o portal F-Tep teria os produtos que melhor se adequavam a estas necessidades, por isso, este foi estudado em maior detalhe, comparativamente aos outros.

No geral notou-se que o portal F-Tep não tem disponíveis produtos que permitam alcançar diretamente os objetivos pretendidos, assim sendo, optou-se por desenvolver metodologias que permitissem aprimorar e/ou gerar resultados qualitativos para a gestão florestal. Estas metodologias foram aplicadas a uma área de estudo, que é apresentada ao longo da dissertação.

Foram gerados vários produtos com o objetivo de identificar zonas de floresta e zonas de alteração de floresta. Dos produtos gerados, uma parte foi validada e mostrou ser competente para a concretização dos objetivos propostos. Os erros apresentados deveram-se ou à recolha dos dados de referência ou à criação dos dados de treino.

Palavras-chave: RAIZ, mapas de ocupação do solo, floresta, *Sentinel 2*, F-Tep.

Abstract

This paper was carried out within the framework of the project “Exploring the Potential of the Sentinel missions Satellite Imagery” (EPSSI) and carried out in collaboration with RAIZ, it was essentially based on the creation of land occupation maps, with a particular focus on the forest. Some tools available on the Platform for Forest Thematic Exploration (F-Tep) were tested for the automatic identification of forest zones and forest alteration areas. All of these objectives were proposed taking into account the importance of forest knowledge in our territory, what type of forests are available, how many and how we can manage their materials in a sustainable way.

Initially, a study of the products available was made for forestry purposes in other entities, in addition to RAIZ, such as: Sentinel Hub, Terramonitor, F-Tep, Giovanni, Land Cover and Global Forest. We can only know if a product will be useful to us when we know its specificities, for that, the characteristics of the products available in each portal were studied: the update period, the cost, the spatial resolution, the nomenclature, the data source and how the products are obtained. After analyzing the existing products in all portals, those that would be useful for an institution like RAIZ were selected. It was noted that the F-Tep portal would have the products that best suited these needs, so it was the one studied in more detail, compared to the others.

In general, it was noted that the F-Tep portal does not have products available to directly achieve the intended objectives, so it was decided to develop methodologies that would allow to improve and / or generate qualitative results for forest management. These methodologies were applied to a study area, which is presented throughout the dissertation.

Several products were generated with the objective of identifying forest zones and areas of forest change. Of the generated products, a part was validated and proved to be competent to achieve the proposed objectives. The errors presented were due either to the collection of reference data or to the creation of training data.

Keywords: RAIZ, land cover maps, forest, Sentinel 2, F-Tep.

Acrónimos, Siglas e Símbolos

Acrónimos e Siglas

BOA – Bottom Of Atmosphere

CLC – Corine Land Cover

LC – Cobertura Foliar

CVF – Cobertura de Vegetação Fracionada

EPSSI – Exploring the Potential of the Sentinel missions Satellite Imagery

ESA – European Space Agency

FN – False Negative

FP – False Positive

F-Tep – Forestry Tep

GEMI – Global Environment Monitoring Index

GPS – Global Positioning System

GRD – Ground Range Detected

IPVI – Índice de Vegetação por Percentagem de Infravermelho

ISP – Instrument Source Packet

ISAUL – Instituto Superior de Agronomia

IUR – Identificador uniforme de recurso

LAI – Índice de Área Foliar

LCLU – Land Cover and Land Use

LCC – Teor de Clorofila na Folha

LWC – Teor de Água nas Folhas

MDE – Modelo Digital de Elevação

MSAVI – Índice de Vegetação Ajustada do Solo Modificado

MSAVI2 – Índice de Vegetação Ajustada do Solo Modificado 2

MSI – Instrumento Multiespectral

NASA – National Aeronautics and Space Administration

NDVI – Normalized Difference Vegetation Index

NBR – Normalized Burn Ratio

NIR – Near-Infrared

NUTS – Nomenclatura de Unidades Territoriais de Estatística

OE – Observação Espacial

OTB – Orfeo Toolbox

QGIS – Quantum Geographic Information System

RAIZ – Instituto de investigação da Floresta e do Papel

RGB – Red, Green and Blue

RVI – Índice de Vegetação de Proporção

S2A – Sentinel 2A

S2B – Sentinel 2B

SAVI – Índice de Vegetação Ajustada do Solo

SCR – Sistema de Coordenadas de Referência

SIG – Sistema de Informação Geográfica

SNAP – Sentinel Application Platform

SWIR – Infravermelho de ondas curtas

TNDVI – Índice de vegetação com diferença normalizada transformada

TOA – Top Of Atmosphere

TSAVI – Índice de Vegetação Ajustada do Solo Transformado

UA – Universidade de Aveiro

UC – Universidade de Coimbra

UE – União Europeia

UMM – Unidade Mínima Mapeável

UOS – Uso e Ocupação do Solo

UTM – Universal Transversal de Mercator

VESA – Vegetation by European Space Agency

VN – Verdadeiro Negativo

VP – Verdadeiro Positivo

VTT – Technical Research Centre of Finland

ZIF – Área de Intervenção Florestal

Símbolos

E_C – Erros de Comissão

E_G – Exatidão Global

E_O – Erros de Omissão

E_P – Exatidão do Produtor

E_U – Exatidão do Utilizador

n_{ii} – Elementos classificados corretamente

n_{+i} – Total de elementos de referência amostrados para uma classe i

n_{i+} – Total de elementos classificados para uma classe i

Índice de Figuras, Tabelas e Fluxogramas

Figuras

Figura 1 – Antigo formato dos nomes dos produtos gerados pelas imagens do <i>Sentinel 2</i> Nível-1C.....	7
Figura 2 – Classificador Random Forest. Fonte: https://medium.com/@ar.ingenious/applying-random-forest-classification-machine-learning-algorithm-from-scratch-with-real-24ff198a1c57	9
Figura 3 - Curvas de resposta espectral em áreas queimadas (linha a vermelho) e em áreas com vegetação saudável (linha a verde). Fonte: http://un-spider.org/advisory-support/recommended-practices/recommended-practice-burn-severity/in-detail/normalized-burn-ratio	11
Figura 4 - Portal <i>Sentinel Hub</i> . Fonte: www.sentinel-hub.com	15
Figura 5 - Mapa de agricultura na zona de Abrantes	16
Figura 6 – Mapa com o índice de vegetação na zona de Abrantes.....	16
Figura 7 - Portal <i>Terramonitor</i> . Fonte: https://www.terramonitor.com/	17
Figura 8 - Portal <i>Giovanni</i> . Fonte: https://giovanni.gsfc.nasa.gov/giovanni/	18
Figura 9 – Mapa com o LAI para Portugal.	19
Figura 10 - Mapa com o LAI para Portugal.	19
Figura 11 - Portal <i>Land Cover</i> . Fonte: https://land.copernicus.eu/global/products/lc	20
Figura 12 - Mapa com todos os tipos de floresta quando selecionada a zona Abrantes.....	21
Figura 13 – Mapa com a floresta aberta e fechada quando selecionada a zona Abrantes.....	21
Figura 14 - Portal <i>Global Forest</i> . Fonte: https://www.globalforestwatch.org/	22
Figura 15 - Mapa da cobertura arbórea na zona de Abrantes.....	23
Figura 16 - Mapa da cobertura do solo na zona de Abrantes.....	23
Figura 17 - Portal F-Tep. Fonte: https://f-tep.com/	24
Figura 18 - Exemplo de <i>output</i> da ferramenta LandCoverS2.....	28
Figura 19 - Exemplo de <i>output</i> da ferramenta VegetationIndices	29
Figura 20 - Exemplo de <i>output</i> da ferramenta ForestChangeS2.....	31
Figura 21 - Exemplo de <i>output</i> da ferramenta CombS2granules	32
Figura 22 - Exemplo de <i>output</i> da ferramenta CombS2granulesZ	33

Figura 23 – Áreas de interesse no território português. Fonte: https://mysustainableforest.com/case-demos/portugal/	38
Figura 24 - Caso de Estudo.	39
Figura 25 - Correspondência entre a nomenclatura CLC e as megaclasses. Fonte: http://mapas.dgterritorio.pt/atom-dgt/pdf-cous/CLC2012/CLC-1990-2000-2006-2012-PT.pdf	41
Figura 26 - Dados de treino da imagem de inverno.....	42
Figura 27 – Separabilidade das classes para a imagem de inverno. (a) Antes da alteração dos polígonos (b) Depois da alteração dos polígonos	43
Figura 28 - Separabilidade das classes para a imagem de outono. (a) Antes da alteração dos polígonos (b) Depois da alteração dos polígonos	43
Figura 29 - Separabilidade das classes para a imagem de verão. (a) Antes da alteração dos polígonos (b) Depois da alteração dos polígonos	44
Figura 30 – Zona problemática, erroneamente classificada como água.....	45
Figura 31 - Mapas de ocupação do solo para as imagens do período de outono (à esquerda) e de verão (à direita).....	46
Figura 32 - Classificação da imagem de verão já com a classe "Queimado".	48
Figura 33 - Classificação da imagem de outono já com a classe "Queimado".	49
Figura 34 - Confusão entre as classes "Solo Nu" e "Queimado" na imagem de outono.....	50
Figura 35 - Confusão entre as classes "Solo Nu" e "Queimado" na imagem de verão.	51
Figura 36 - Classificação final da imagem de outono.....	52
Figura 37 - Classificação final da imagem de verão.	52
Figura 38 – Máscara da floresta para o período de inverno.	53
Figura 39 - Máscara da floresta para o período de outono.....	54
Figura 40 - Máscara da floresta para o período de verão.	54
Figura 41 - Pontos perto de caminhos rurais. Na primeira linha de imagens temos como base layer o basemap do ArcMap, na segunda linha temos como base layer a classificação de verão, sendo que a verde claro temos a classe “Vegetação herbácea e arbustiva”, a verde	56
Figura 42 - Pontos em plantações jovens. Na primeira linha de imagens temos como base layer o basemap do ArcMap, na segunda linha temos como base layer a classificação de verão, sendo que a verde claro temos a classe “Vegetação herbácea e arbustiva”, a verde escuro a classe “Floresta” e a castanho claro a classe “Solo Nu”.	57
Figura 43 - Pontos em floresta dispersa e com vegetação e solo nu. Na primeira linha de imagens temos como base layer o basemap do ArcMap, na segunda linha temos como base layer a classificação de verão, sendo que a verde claro temos a classe “Vegetação herbácea e arbustiva”, a verde escuro a classe “Floresta” e a castanho claro a classe “Solo Nu”.	57
Figura 44 - Zonas de alteração entre o verão e o inverno. Em cima a imagem de verão seguida das alterações e por baixo a imagem de inverno seguida das alterações.....	61
Figura 45 - Zonas de alteração entre o outono e o inverno. Em cima a imagem de inverno seguida das alterações e por baixo a imagem de outono seguida das alterações.....	61

Figura 46 - Zonas de alteração entre o verão e o outono. Em cima a imagem de verão seguida das alterações e por baixo a imagem de outono seguida das alterações.62

Tabelas

Tabela 1 - Matriz de Confusão	13
Tabela 2 - Quadro geral das plataformas estudadas.	26
Tabela 3 - Matriz de confusão para a classificação de inverno.	55
Tabela 4 - Matriz de confusão para a classificação de verão.....	56

Fluxogramas

Fluxograma 1 - Síntese metodológica dos passos desenvolvidos na parte prática	37
Fluxograma 2 – Processo para a classificação de mapas com zonas queimadas.....	46
Fluxograma 3 – Processo para a escolha do melhor método para a identificação de zonas de alteração na floresta.....	58

1. Introdução

O solo é a camada superior da crosta, entre a litosfera e a atmosfera (Ríos et al., 2018). Segundo Hedley et al. (2018), o uso do solo é a forma pela qual o espaço geográfico está sendo utilizado por parte do ser humano e a ocupação do solo diz respeito a como o mesmo está ocupado, ou seja, os materiais físicos existentes sobre este. São dois termos (Uso e Ocupação do Solo – UOS) que não podem ser vistos como independentes um do outro. O UOS representa a interação homem-ambiente dentro de determinada área. Os mapas de UOS são utilizados em várias aplicações, tais como, na representação espacial de ocupação do solo, no cálculo da área das diferentes classes, na avaliação de alterações, no planeamento e na gestão do território, entre outros (Stehman, 2009).

Os processos de mapeamento podem ser realizados pela execução de diferentes métodos, de baixo ou alto custo, menos ou mais sofisticados. Tradicionalmente, para se obter mapas de qualidade é necessário um levantamento direto do solo, proporcionando uma maior precisão nos resultados, no entanto, são métodos muito dispendiosos em tempo e recursos financeiros. Uma abordagem alternativa à tradicional é recorrer ao levantamento de dados do terreno por fotointerpretação, facilitando o planeamento, a execução e a monitorização da atividade para a gestão florestal, proporcionando menores custos e um menor impacto ambiental. Embora este método exija menos tempo do que o tradicional, por consequência pode exigir um alto custo na obtenção de imagens de alta resolução (Cenggoro et al., 2016). No entanto, nem em todos os trabalhos se justifica a utilização de imagens de alta resolução e, nesse caso, são utilizadas imagens de mais baixa resolução em que o acesso é livre e sem custos, como é o caso das imagens do *Sentinel 2*.

Devido às limitações dos processos manuais, a Detecção Remota tornou-se bastante popular no mapeamento do solo. Segundo Cenggoro et al. (2016), a sua popularidade deve-se às amplas aplicações nas diferentes áreas, relativamente à área da floresta, temos como exemplos a monitorização: de alterações na floresta, da ocupação e uso do solo, de queimadas, de desflorestação e entre outros. Deve-se também à automação de processos e à facilidade de acesso das imagens obtidas, o seu uso levou ao desenvolvimento de várias plataformas como as que são abordadas ao longo da dissertação.

Em Portugal ocorreram grandes transições de UOS nas últimas duas décadas, houve uma redução de área florestal e de alguns tipos de áreas agrícolas, e deu-se, um aumento das áreas urbanas (Meneses et al., 2018). A análise da floresta em Portugal é muito complexa, devido à litologia, à topografia, aos solos (devido à diversidade das rochas matriz, termo mais conhecido como rocha-mãe) e ao clima.

A floresta deve ser planeada e estruturada de acordo com os interesses ambientais e económicos do país, sem legitimar eventuais impactos ambientais criados (Garcia, 2017).

1.1 Enquadramento e motivação

O projeto realizado foi enquadrado no projeto EPSSI (financiado pelo Instituto de Engenharia de Sistemas e Computadores de Coimbra – INESC Coimbra), que tem como objetivo explorar o potencial do uso de imagens de satélite recolhidas pelas missões *Sentinel* do programa *Copernicus* da *European Space Agency* (ESA) para as necessidades da sociedade contemporânea (INESC Coimbra, 2020), em colaboração com o centro de investigação RAIZ.

O RAIZ é um organismo privado que tem como sócios a *The Navigator Company*, a Universidade de Aveiro (UA), a Universidade de Coimbra (UC) e o Instituto Superior de Agronomia da Universidade de Lisboa (ISAUL). Foi criado em 1996, com a missão de contribuir para a competitividade e sustentabilidade da fileira do eucalipto, da floresta ao produto: “gerando conhecimento de excelência, prestando serviços e consultoria especializados, desenvolvendo soluções inovadoras promotoras de uma nova bioeconomia, promovendo a cooperação com universidades e centros de conhecimento de classe mundial e formando recursos humanos altamente qualificados e motivados.” (RAIZ, 2020). Tem como visão “Ser reconhecido a nível mundial como um centro de investigação de referência, promotor do desenvolvimento sustentável e da bioeconomia baseada na floresta do eucalipto” (RAIZ, 2020).

Relativamente ao UOS e aos parâmetros de floresta, o RAIZ tem como critério a manutenção e o aprimoramento dos recursos florestais e a sua contribuição para os ciclos globais de carbono (RAIZ, 2020). Atualmente, são utilizados métodos de amostragem e técnicas de medição de árvores, estes processos geralmente consomem tempo, são caros e as informações recolhidas não são precisas e ficam desatualizadas rapidamente. O portal do RAIZ tem como produtos:

- Uma máscara da floresta,
- A delimitação da massa florestal,
- As infraestruturas florestais,
- Os principais tipos de floresta,
- A idade da floresta,
- As cicatrizes de queimadura,
- Os cortes claros,
- O DEM, a altitude, o declive e o aspeto,
- A biomassa acima do solo,
- O stock de dióxido de carbono.

O mapeamento de florestas é bastante importante para podermos ter noção de como está a ser usado e ocupado o solo, que tipo de materiais têm disponíveis, que tipo de espécies estão presentes e em que áreas estas são predominantes. Conhecer o território de uma nação é um passo crucial para se conseguir a manutenção da segurança nacional e para contribuir com o seu desenvolvimento (de Freitas Borges & Lucas Rajão, 2016).

A motivação para este projeto tem enfoque na floresta e na contribuição para a identificação de fontes de dados que possam ser úteis a uma instituição como o RAIZ.

1.2 Objetivos

Numa fase inicial de recolha e análise de dados, este projeto tem como objetivos:

- 1) Identificar plataformas que disponibilizam informação sobre a cobertura do solo, potencialmente úteis para a monitorização da floresta,
- 2) Analisar a informação disponível nas várias plataformas e as suas características,
- 3) Selecionar a ou as plataformas mais promissoras e fazer uma análise dos produtos disponíveis
- 4) Explorar as ferramentas disponibilizadas para obter informação sobre a identificação das zonas de floresta e alterações sofridas.

Com base nas plataformas que oferecem serviços para a monitorização florestal, na análise e no teste dos produtos disponíveis, dos benefícios e das limitações, de entre todas as necessidades de informação do RAIZ foi decidido focar este projeto essencialmente em dois objetivos:

- 1) Identificação automática do que é floresta e não floresta,
- 2) Identificação automática de zonas de alteração da floresta.

Como objetivos secundários temos a criação de mapas de ocupação do solo, a validação da máscara de floresta para o período de inverno e verão e a validação da máscara com as alterações da floresta.

1.3 Organização da dissertação

A dissertação está organizada por capítulos, que passam por abordar questões teóricas até à realização prática das mesmas. No primeiro capítulo, é feita uma introdução do que será realizado, os motivos para a sua realização e os respetivos objetivos da dissertação. No segundo capítulo, são apresentados conceitos base de deteção remota para o entendimento de posteriores secções. No terceiro capítulo, é feita uma análise de portais existentes para a monitorização florestal, são analisados os produtos com potencial utilização por institutos como o RAIZ e é selecionado aquele

que foi considerado como mais adequado para as necessidades do projeto. No quarto capítulo é explorado esse portal em maior detalhe, assim como as ferramentas que este dispõe. No quinto capítulo é descrito todo o procedimento prático para a identificação automática de zonas de floresta e de alteração de floresta, onde é utilizado o portal escolhido anteriormente. Por fim, no sexto capítulo, são discutidos os resultados obtidos pelos produtos gerados e são tiradas conclusões sobre os mesmos.

2. Conceitos de Detecção Remota

Para uma boa compreensão da dissertação é necessário ser feita uma introdução a alguns conceitos base no ramo da detecção remota. Foram utilizadas imagens recolhidas pelos satélites das missões *Sentinel*, pelo que o programa *Sentinel 2* é apresentado na secção 2.1. Para fazer a classificação de imagens de satélite com classificadores supervisionados é necessário definir áreas de treino, estas áreas são explicadas na secção 2.2. É também abordada a classificação de imagens e o algoritmo utilizado para a classificação das mesmas. Podem ser calculados diversos índices radiométricos que ajudam na classificação de imagens, a partir das bandas de imagens de satélite, nomeadamente o *Normalized Burn Ratio* (NBR) e o *Normalized Difference Vegetation Index* (NDVI), como é exemplificado na secção 2.3. Por fim, na secção 2.4 é descrito o processo de avaliação da exatidão utilizado.

2.1. A missão *Sentinel 2*

Os satélites *Sentinel* foram desenvolvidos no âmbito da Componente Espaço do programa *Copernicus*, que tem como objetivo a recolha, o pré-processamento e a distribuição de dados obtidos a partir de sensores remotos instalados em satélites.

Os satélites existentes e previstos para um futuro próximo são: o *Sentinel 1*, o *Sentinel 2*, o *Sentinel 3*, o *Sentinel 4*, o *Sentinel 5*, o *Sentinel 6* e o *Sentinel 7*. Como o *Sentinel 2* é dedicado essencialmente à monitorização do meio terrestre, e uma vez que o projeto é direcionado à monitorização florestal, as informações dos dados obtidos pelo satélite *Sentinel 2* serão especificadas mais detalhadamente. O satélite *Sentinel 1* também apresenta informação útil para a monitorização florestal, no entanto, as imagens recolhidas por estes satélites não foram utilizadas neste projeto e assim sendo, esta componente não foi desenvolvida neste trabalho.

O *Sentinel 2* é uma missão europeia que recolhe imagens multiespetrais de espectro amplo e de alta resolução (Delwart, 2015). O programa *Sentinel 2* é constituído por um par de satélites, o satélite 2A (S2A) e o satélite 2B (S2B), que quando operam em simultâneo permitem uma maior frequência na aquisição de imagens. As imagens recolhidas têm alta resolução espacial e temporal (5 dias no equador, se os dois satélites forem usados em simultâneo e 10 dias, se só for usado um deles).

Segundo Delwart (2015), os dados de observação adquiridos na missão *Sentinel 2* são utilizados por serviços como:

- Monitorização do solo – Fornecendo dados de entrada para o mapeamento da mudança da cobertura do solo e parâmetros biofísicos (como o Índice de Área de Folha – LAI, o Teor de Clorofila da Folha – LCC, a Cobertura Foliar – LC);
- Gestão de emergências – Fornecendo dados de observação por satélites de forma rápida e eficiente, para apoiar a gestão dos esforços de assistência a desastres naturais, para conhecer os edifícios ameaçados e as estruturas complexas que foram previamente identificadas como estando em risco de desastres naturais;
- Segurança – Em questões como a vigilância de fronteiras e a vigilância marítima, para apoio à ação externa da União Europeia (UE);
- Mudanças climáticas.

O instrumento multiespectral (MSI) a bordo dos satélites do *Sentinel 2* trabalha passivamente recolhendo a luz solar refletida na Terra. Novos dados são adquiridos no instrumento à medida que o satélite se movimenta (Delwart, 2015). Segundo Hedley et al. (2018), o instrumento multiespectral a bordo do *Sentinel 2* gera 13 bandas espectrais que são concebidas por dois sensores:

- 4 bandas no visível e no infravermelho,
- 6 bandas no “*red edge*” e no infravermelho de ondas curtas,
- 3 bandas para correções atmosféricas.

A órbita do *Sentinel 2* é síncrona ao sol, isto é, realiza-se em simultâneo com o sol. Este tipo de órbita é utilizado para garantir que o ângulo da luz solar na superfície da Terra é mantido constante, o que permite minimizar o impacto das sombras (Delwart, 2015). Isto garante uma consistência ao longo do tempo, sendo isso crítico para a avaliação de dados de séries temporais.

Em cobertura geográfica, os satélites *Sentinel 2* fornecem dados sobre áreas terrestres e costeiras numa faixa de latitude que se estende de 56° sul a 83° norte (Delwart, 2015). A missão *Sentinel 2* tem uma cobertura geográfica de todas as áreas terrestres e ilhas, exceto da Antártica. Os dados apresentam uma unidade mínima cartográfica (UMM) inferior a 1ha, isto é, 1ha é a tamanho do menor elemento que está a ser mapeado, de forma confiável no mapa.

Segundo Delwart (2015), os principais produtos disponíveis no *Sentinel 2* são:

- O Nível-0, que não está disponível para os utilizadores, é composto por dados de imagem bruta compactados no formato *Instrument Source Packet* (ISP) e contém todas as informações necessárias para gerar o Nível 1 ou superior.
- O Nível-1A, que não está disponível para os utilizadores, é obtido descompactando os dados da imagem bruta de “Nível-0”.
- O Nível-1B, que é o nível mais baixo disponível para os utilizadores, fornece imagens corrigidas radiometricamente para valores de refletância da *Top Of*

Atmosphere (TOA). Os valores de refletância TOA representam a refletância bruta da Terra medida do espaço, isto é, são uma mistura da luz refletida na superfície da Terra e na atmosfera.

- O Nível-1C resulta do uso de um Modelo Digital de Elevação (MDE) para projetar a imagem em coordenadas cartográficas, as coordenadas do pixel referem-se ao canto superior esquerdo do pixel. Fornece ao utilizador imagens ortorretificadas;
- O Nível-2A que fornece imagens de refletância da *Bottom Of Atmosphere* (BOA), sendo estas derivadas dos produtos Nível-1C e associados. Os valores de refletância BOA representam a refletância real das áreas na superfície da Terra.

Em 6 de Dezembro de 2016 foi alterado o formato de convenção dos nomes de produtos gerados pelas imagens do *Sentinel 2* Nível-1C. O antigo formato e a explicação de cada componente é o ilustrado na **Figura 1**.

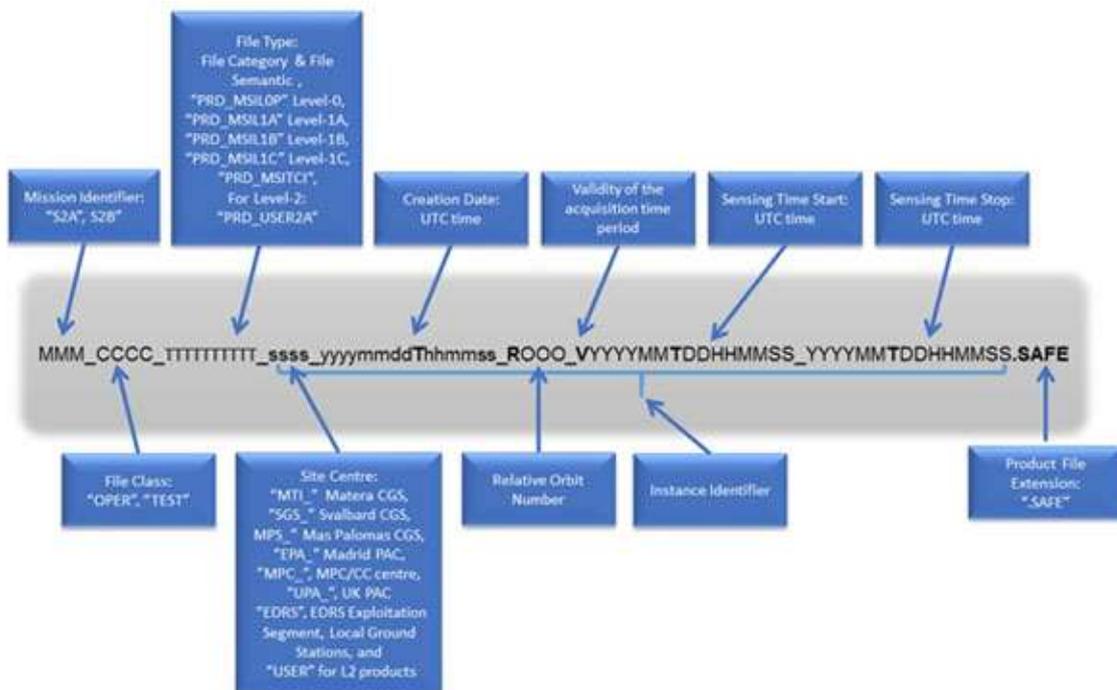


Figura 1 – Antigo formato dos nomes dos produtos gerados pelas imagens do *Sentinel 2* Nível-1C.

Atualmente o formato de convenção dos nomes dos produtos gerados pelas imagens obtidas pelo *Sentinel 2* Nível-1C é organizado do seguinte modo:

- MMM_MSIXXX_YYYYMMDDHHMMSS_Nxxyy_ROOO_Txxxxx_YYYYMM DDHHMMSS.SAFE

Tendo em conta que:

- 1) MMM: ID da missão,

- 2) MSIXXX: Nível do sensor multiespectral,
- 3) YYYYMMDDHHMMSS: Data do início e horas da detecção do sensor,
- 4) Nxyyy: Número da linha de base de processamento,
- 5) ROOO: Número de órbita relativa,
- 6) Txxxxx: ID da área de pesquisa,
- 7) YYYYMMDDHHMMSS: Data do fim e horas da detecção do sensor,
- 8) SAFE: Formato do produto.

Segundo Gatti & Naud (2017), o formato atual foi introduzido para superar a limitação de 256 caracteres nos nomes de caminho impostos pelas plataformas *windows*, para isso foram compactados os nomes dos arquivos, incluindo a nomenclatura de pastas e arquivos internos.

2.2. Classificação de imagens multiespectrais

Para a classificação de imagens multiespectrais podem ser utilizados classificadores supervisionados e não supervisionados. Nos classificadores supervisionados é o utilizador que define *à priori* as classes espectrais e que seleciona os dados de treino, enquanto que nos classificadores não supervisionados não há a intervenção do utilizador, os pixéis são atribuídos às classes espectrais sem o conhecimento prévio do mesmo (Torres-Moreno et al., 2009).

A classificação supervisionada é muito utilizada devido a ser um processo mais rápido que o método não supervisionado, principalmente quando se trabalha com grandes áreas (Crósta, 1992). Por isso, neste projeto foi utilizado um classificador supervisionado, e assim sendo, foi necessária a criação dos dados de treino, que consistem num conjunto de amostras. Uma amostra é uma parte da região mapeada e permite selecionar locais nos quais os dados de referência serão recolhidos (Stehman, 2009).

Para a criação dos dados de treino tem de ser definido *à priori* a nomenclatura que vai ser utilizada para classificar a imagem. Depois de se saber que classes serão consideradas, são criados os polígonos para cada classe por fotointerpretação de imagens de alta resolução. Devem ser criados tantos polígonos quanto os necessários para uma boa caracterização das classes. Os polígonos criados devem estar distribuídos uniformemente pela área de estudo e devem ser representativos das várias classes.

O algoritmo supervisionado utilizado neste projeto para a classificação da cobertura do solo é o *Random Forest* (**Figura 2**). Este algoritmo baseia-se nas amostras das classes de cobertura do solo, com estas é produzido um conjunto de árvores de decisão, na **Figura 2** este conjunto corresponde à “TREE #1”, “TREE #2”, “TREE #3” e “TREE #4”. Para a obtenção desse conjunto, é feita a seleção de um subconjunto aleatório e a reposição do mesmo. Como desvantagem, uma

vez que é feita a reposição pode acontecer a seleção do mesmo conjunto de dados uma vez, duas ou mais vezes, podendo até não haver a seleção de certos conjuntos. Em seguida, é criado um modelo de decisão em árvore com regras binárias, que tem como objetivo o encontro das classes que pretendemos estudar, na **Figura 2** estas classe correspondem à “CLASS B”, “CLASS C” e “CLASS D”. Para cada árvore, é escolhido um conjunto aleatório de bandas espectrais em cada nó de decisão. A classe final de cobertura do solo é escolhida com base na votação maioritária, isto é, é escolhida a classe com maior percentagem de ocorrência entre os resultados de cada árvore de decisão, no caso representado pela **Figura 2** a classe final seria a “CLASS C”. Este algoritmo tem como vantagem uma boa adaptabilidade a diferentes problemas de classificação.

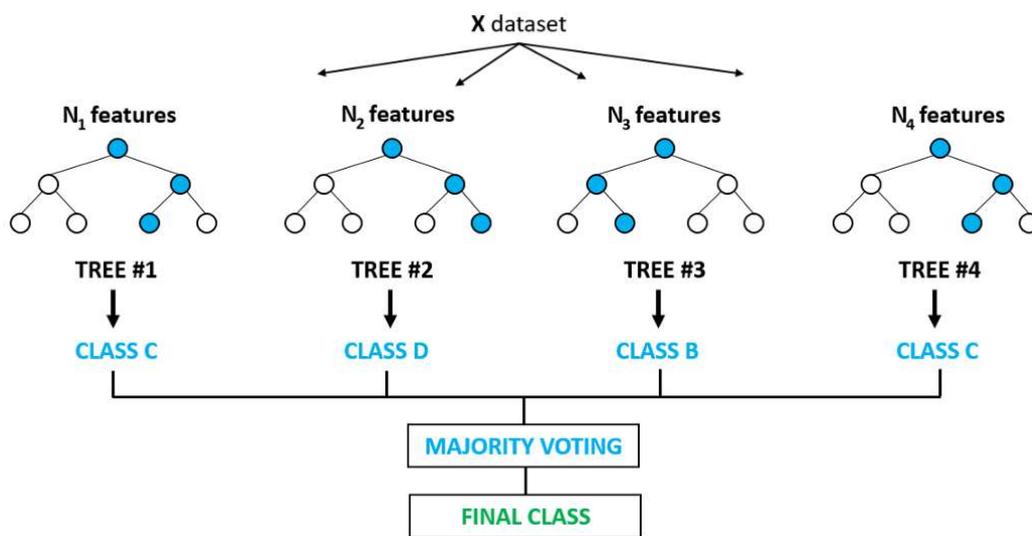


Figura 2 – Classificador Random Forest. Fonte: <https://medium.com/@ar.ingenious/applying-random-forest-classification-machine-learning-algorithm-from-scratch-with-real-24ff198a1c57>

2.3. Índices radiométricos

Em radiometria é medida a energia contida na radiação eletromagnética em todos os comprimentos de onda. Existem diversos índices radiométricos, como são exemplo:

- NDVI

O NDVI é o índice mais utilizado para a identificação de áreas com vegetação e as suas condições (Jinru & Baofeng, 2017). Este índice aproveita o contraste de características entre duas bandas de um conjunto de dados multiespectral *raster*, a absorção do pigmento de clorofila na banda vermelha (*Red*) e a alta refletividade do material vegetal na banda do infravermelho próximo (NIR). É obtido pela **Equação 1**:

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (1)$$

Segundo Jinru & Baofeng (2017), uma vez que o NDVI é calculado por meio de um procedimento de normalização, a faixa de valores está entre os valores -1 e 1. O valor obtido representa a presença de vegetação, quanto maior ele é, maior é quantidade vegetativa do local. Valores negativos ou próximos a zero indicam áreas de água ou solo nu, onde a clorofila é reduzida e com isso há baixa quantidade de vegetação.

No entanto, o NDVI é sensível aos efeitos do brilho do solo, cor do solo, atmosfera, nuvens e sombra de nuvens, entre outros (Jinru & Baofeng, 2017). Portanto, uma série de derivados e alternativas ao NDVI foram propostas para abordar essas limitações:

- Transformed Normalized Difference Vegetation Index (TNDVI)

O TNDVI indica uma correlação entre a quantidade de biomassa verde e a que é encontrada num pixel (Bannari et al., 2002).

- Ratio Vegetation Index (RVI)

O RVI fornece uma indicação geral da vegetação com base que as folhas absorvem relativamente mais luz vermelha do que luz infravermelha (Jinru & Baofeng, 2017). Valores baixos indicam solo, água ou gelo e valores altos indicam vegetação saudável (Driese, 2012).

- Soil Adjusted Vegetation Index (SAVI)

O SAVI é um índice de vegetação que tenta minimizar as influências do brilho do solo, usando um fator de correção para o brilho (Driese, 2012).

- Transformed Soil Adjusted Vegetation Index (TSAVI)

O TSAVI é um índice de vegetação que tenta minimizar as influências do brilho do solo, assumindo que a linha do solo tem um declínio e uma interceção arbitrária (Driese, 2012).

- Modified Soil Adjusted Vegetation Index (MSAVI)

O MSAVI é um índice de vegetação que tenta minimizar o efeito do solo nu no SAVI (Jinru & Baofeng, 2017).

- Modified Soil Adjusted Vegetation Index 2 (MSAVI2)

O MSAVI2 é um índice de vegetação que elimina a necessidade de encontrar a linha do solo ou até mesmo especificar explicitamente o fator de correção de brilho do solo (Jinru & Baofeng, 2017).

- Global Environment Monitoring Index (GEMI)

O GEMI é um índice de vegetação desenvolvido para minimizar os efeitos atmosféricos e do solo simultaneamente, não sendo recomendado o seu uso em zonas áridas (Driese, 2012).

- Infrared Percentage Vegetation Index (IPVI)

O IPVI é funcionalmente equivalente ao NDVI e ao RVI mas, apenas varia entre 0 e 1, e é computacionalmente mais rápido (Driese, 2012).

- NBR

O NBR foi desenvolvido para destacar áreas queimadas e estimar a severidade do fogo (Escuin et al., 2008). A fórmula é semelhante à do NDVI, no entanto, usa comprimentos de onda do infravermelho de ondas curtas (SWIR). Tendo em conta as características relativas dos vários materiais e as bandas do sensor *Sentinel 2*, para o estudo de áreas ardidas são usadas as bandas 8 e 12. A fórmula é expressa pela **Equação 2**:

$$\text{NBR} = \frac{\text{NIR} - \text{SWIR}}{(\text{NIR} + \text{SWIR})} \Leftrightarrow \text{NBR} = \frac{\text{Banda 8} - \text{Banda 12}}{\text{Banda 8} + \text{Banda 12}} \quad (2)$$

Segundo Escuin et al. (2008), na NIR a vegetação saudável apresenta valores altos de refletância e valores baixos na SWIR (como pode ser visto na **Figura 3**). O contrário ocorre nas áreas queimadas, baixos valores de refletância na NIR e altos valores na SWIR (como pode ser visto na **Figura 3**). Com isto, um alto valor de NBR indica que estamos perante uma área de vegetação saudável, enquanto que um baixo valor indica que estamos perante uma zona sem vegetação (solo descoberto) ou uma área recentemente ardida.

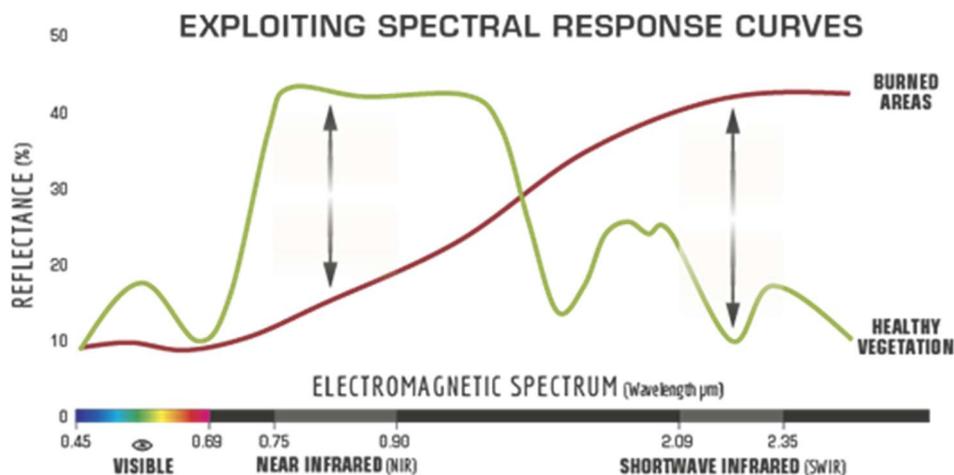


Figura 3 - Curvas de resposta espectral em áreas queimadas (linha a vermelho) e em áreas com vegetação saudável (linha a verde). Fonte: <http://un-spider.org/advisory-support/recommended-practices/recommended-practice-burn-severity/in-detail/normalized-burn-ratio>

2.4. Validação de mapas de ocupação do solo

Os mapas devem ser sempre validados, a validação dos mesmos tem particular importância quando são usados para tomar decisões em várias áreas, como para fins científicos e de gestão de recursos. Nenhum projeto de classificação da cobertura de solo está completo sem uma avaliação da exatidão (Cihlar 2000).

Segundo Stehman (2009), a exatidão de uma classificação de uso e cobertura do solo é o grau em que o mapa a validar concorda com a classificação de referência que vai ser usada para o validar. Baseia-se na comparação da cobertura do solo representada no mapa com a verdadeira condição de cobertura do solo. Obter a “verdade do terreno” não é normalmente alcançável na prática, por isso as avaliações da exatidão avaliam o mapa da cobertura do solo em relação a um mapa com qualidade superior. No entanto, dados de qualidade superior são muito caros e difíceis de obter para toda a região de interesse. A amostragem estatística torna-se assim uma metodologia valiosa e crítica para a avaliação da exatidão.

Segundo Stehman (2009), os estratos são grupos de pixéis construídos de modo que cada pixel pertença a exatamente um estrato e os estratos formam uma partição da população de todos os pixéis. A amostragem aleatória estratificada por classe consistiu, inicialmente, na escolha de quantos pontos ou regiões (grupos de pixéis) por classe iriam ser considerados, para a posterior validação dos produtos gerados. Cada classe teve o mesmo peso, ou seja, não existiram classes com mais pontos ou mais regiões que outras. A seleção das amostras foi feita aleatoriamente. Com isto, é obtido o conjunto dos dados de referência.

A comparação entre os dados obtidos e os dados de referência pode ser feita por meio de uma matriz de confusão, que relaciona os valores do mapa com os valores reais. A matriz que é demonstrada na **Tabela 1** considera apenas as classes 0 e 1, representando o 0 a ausência da classe e o 1 a presença.

Neste caso (**Tabela 1**), a matriz de confusão apresenta nas linhas os dados de referência e nas colunas os dados do mapa classificado (ou previstos). Os verdadeiros positivos (VP) dizem respeito à classe que procuramos prever e que foi corretamente classificada. Os falsos positivos (FP) dizem respeito à classe que procuramos prever e que foi erroneamente classificada. Os verdadeiros negativos (VN) dizem respeito à classe que não procuramos prever e que foi corretamente classificada. Os falsos negativos (FN) dizem respeito à classe que não procuramos prever e que foi erroneamente classificada.

		Valores Previstos	
		0	1
Valores Reais	0	VN	FP
	1	FN	VP

Tabela 1 - Matriz de Confusão

Uma matriz de confusão exibe a percentagem de pontos que, para uma amostra considerada, é corretamente e incorretamente classificada para os diferentes tipos de cobertura de solo (Stehman, 2009).

Os elementos da diagonal correspondem a classes corretamente classificadas e os que estão fora da diagonal correspondem às classes mal classificadas (erros de classificação: comissão e omissão) (Stehman, 2009). Os elementos presentes nas colunas dizem respeito ao utilizador e os presentes nas linhas ao produtor.

A partir dos erros de omissão (E_O), falta de amostras numa classe à qual na realidade pertencem, e dos erros de comissão (E_C), inclusão de amostras numa classe à qual na realidade não pertencem, são calculadas as exatidões do produtor (E_P) e do utilizador (E_U).

A E_P é calculada a partir dos dados de referência, dividindo o número total de amostras corretamente classificadas numa determinada classe pelo total da linha. A fórmula é apresentada na **Equação 3**.

$$E_P = \frac{n_{ii}}{n_{+i}} \quad (3)$$

Onde:

E_P : exatidão do produtor;

n_{ii} : elementos classificados corretamente;

n_{+i} : total de elementos de referência amostrados para uma classe i .

A E_U é calculada a partir dos dados da classificação, dividindo o número total de amostras corretamente classificadas numa classe pelo total da coluna. A fórmula é apresentada na **Equação 4**.

$$E_U = \frac{n_{ii}}{n_{i+}} \quad (4)$$

Onde:

E_P : exatidão do utilizador;

n_{ii} : elementos classificados corretamente;

n_{+i} : total de elementos classificados para uma classe i .

A exatidão global (E_G) trata-se da proporção de resultados corretos que o classificador alcançou, ou seja, a razão entre as predições corretas pelo total de amostras. É calculada segundo a **Equação 5**.

$$E_G = \frac{TP + TN}{TP + FP + TN + FN} \quad (5)$$

3. Portais de monitorização Florestal

Atualmente, já existem vários produtos e ferramentas disponíveis com o objetivo de dar apoio à gestão e monitorização da floresta. Neste capítulo são apresentados os portais analisados e alguns dos produtos e ferramentas disponíveis, bem como as várias informações e as suas especificações técnicas.

3.1. Sentinel Hub

O portal *Sentinel Hub* (**Figura 4**) permite o processamento de *petabytes* de dados de satélite e facilita o acesso à navegação, visualização e análise das várias imagens de Observação Espacial (OE) (Sentinel Hub, 2020). Apresenta uma interface intuitiva e amigável, sem complicações.

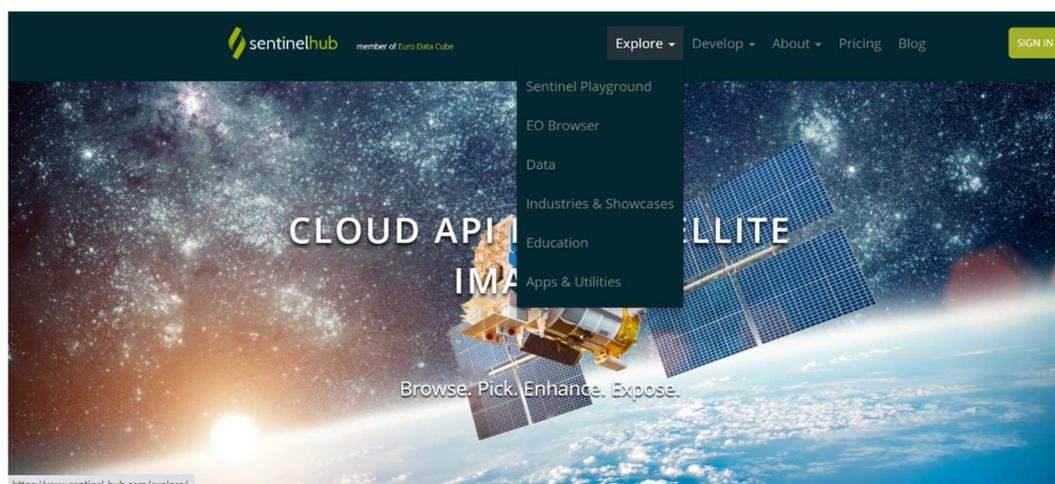


Figura 4 - Portal *Sentinel Hub*. Fonte: www.sentinel-hub.com

Para a exploração de imagens e produtos, este portal subdivide-se em *Sentinel Playground* e *Earth Observation Browser (EO Browser)*. O *EO Browser* torna possível a navegação e a comparação de imagens de todas as fontes de dados que o portal fornece (*Sentinel 1*, *Sentinel 2*, *Sentinel 3*, *Sentinel 5P*, *Landsat*, *Envisat Meris*, *MODIS*, *Proba-V* e *GIBS*), o utilizador escolhe a área de interesse, seleciona o intervalo de tempo das imagens a utilizar e a percentagem de cobertura de nuvens desejada, e com isto, são obtidos os dados no navegador. O *Sentinel Playground* permite a fácil descoberta e exploração das várias imagens do *Sentinel 1*, *Sentinel 2*, *Landsat 8*, *DEM* e *MODIS*, junto com o acesso aos produtos de dados de OE.

Os produtos disponíveis de interesse para este projeto, estão inseridos no *Sentinel Playground* e são estes: os mapas de agricultura (baseados nas bandas 11, 8 e 2 do satélite *Sentinel 2* L1-C) e ao mapas com o índice de vegetação (baseados na combinação de bandas: $\frac{\text{banda 8} - \text{banda 4}}{\text{banda 8} + \text{banda 4}}$)

do satélite *Sentinel 2* L1-C), os dois produtos apresentam uma resolução espacial máxima de 100m e o período de atualização destes é de 5 dias, se os dois satélites S2A e S2B trabalharem em simultâneo, ou 10 dias se só for utilizado um deles.

Na **Figura 5** e na **Figura 6** são apresentados exemplos de *outputs* dos produtos: mapas de agricultura e mapas com o índice de vegetação, respetivamente. Os produtos apresentados estão localizados na zona de Abrantes, apresentam uma resolução espacial de 500m.



Figura 5 - Mapa de agricultura na zona de Abrantes

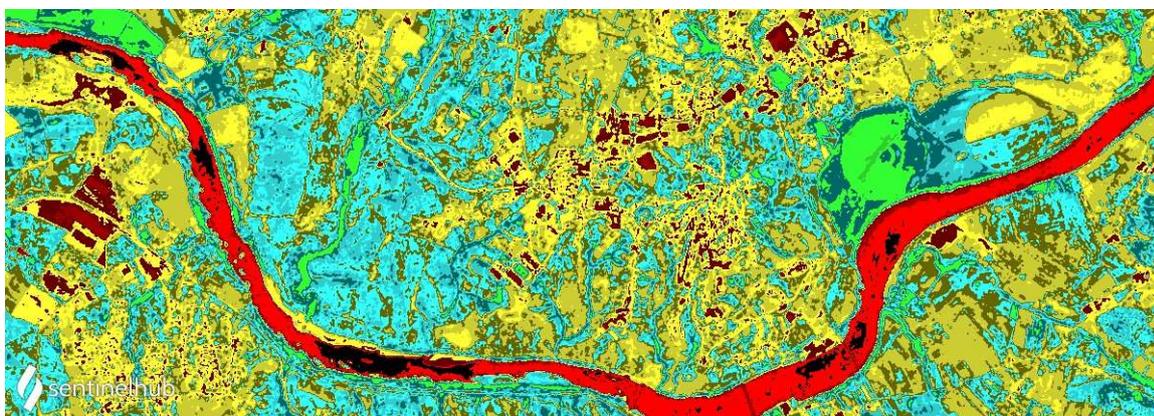


Figura 6 – Mapa com o índice de vegetação na zona de Abrantes.

3.2. *Terramonitor*

O *Terramonitor* (**Figura 7**) é um portal de mapeamento, abrangente e livre de nuvens, que consiste em mais de 100 milhões de imagens (Terramonitor, 2020). O portal permite aceder a dados dos vários satélites *Sentinel*, dependendo do nível de *zoom* e da extensão dos dados. A resolução espacial do mapa base depende da extensão dos dados de satélite e varia entre 10 metros a 1 centímetro por pixel, sendo que as imagens de muito alta resolução são serviços pagos. A resolução temporal varia consoante os produtos, podendo ser de um ou dois dias ou de 5 em 5 dias.

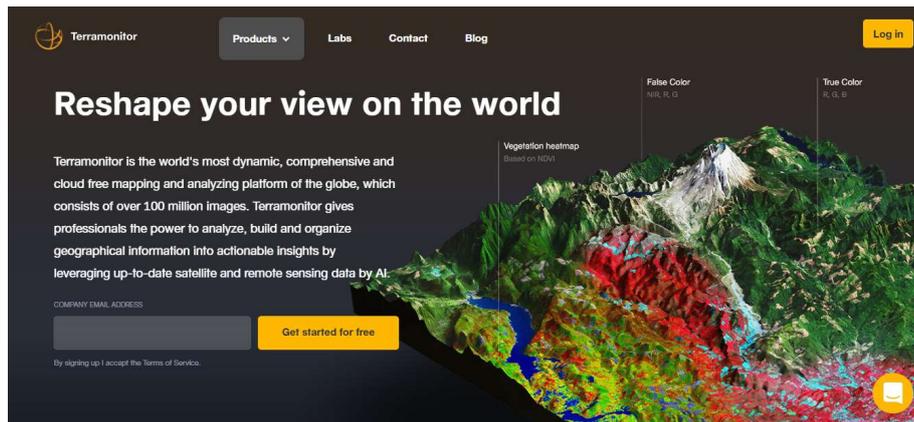


Figura 7 - Portal *Terramonitor*. Fonte: <https://www.terramonitor.com/>

Esta plataforma desenvolveu um serviço de mapeamento baseado em *machine learning*, o índice *Vegetation by European Space Agency* (VESA), que “permite calcular a probabilidade da existência de arbustos prejudiciais em zonas de cultivo de plantas novas” (Terramonitor, 2020). É calculado usando os dados do satélite do S2A e S2B e é classificado em cinco classes: muito alto, alto, moderado, baixo e muito baixo.

Todos os mapas no *ForestMonitor* (serviço da *Terramonitor* desenvolvido para proprietários de florestas obterem imagens de satélite mais atualizadas de florestas e outras áreas de solo) não apresentam nuvens nem neve e recebem constantemente novos dados de satélites (Terramonitor, 2020).

Os produtos disponíveis de interesse para este projeto, existentes no portal são: mapas de vegetação, mapas com o índice VESA e os mapas do serviço *ForestMonitor*. Os produtos apresentam uma resolução espacial de 5m, 10m e 10m, respetivamente. Estes produtos são gerados a partir de imagens obtidas pelo satélite *Sentinel 2*. O período de atualização para os mapas de vegetação e para o produto *ForestMonitor* é de 1 a 2 dias e para o índice VESA é de 5 em 5 dias. O acesso a estes produtos é pago e por isso, não serão apresentados exemplos de *outputs* dos produtos.

O portal *Terramonitor* está a implementar serviços de processamento de dados no ambiente F-Tep, estes serviços são descritos na secção 3.6.

3.3. *Giovanni*

O *Giovanni* (**Figura 8**) é um portal que exhibe dados da Terra dos satélites da *National Aeronautics and Space Administration* (NASA) diretamente na *internet*, sem dificuldade na aquisição e análise de dados. Os dados do portal podem ser filtrados: pelo tipo de observação, pela disciplina de interesse, pela medida que se pretende calcular, pela fonte de dados, pela resolução

espacial e temporal, pelo comprimento de onda, pela profundidade, pelas *features* especiais (referentes à climatologia) e pelo portal. Para cada um destes filtros são apresentados os mais diversos resultados.

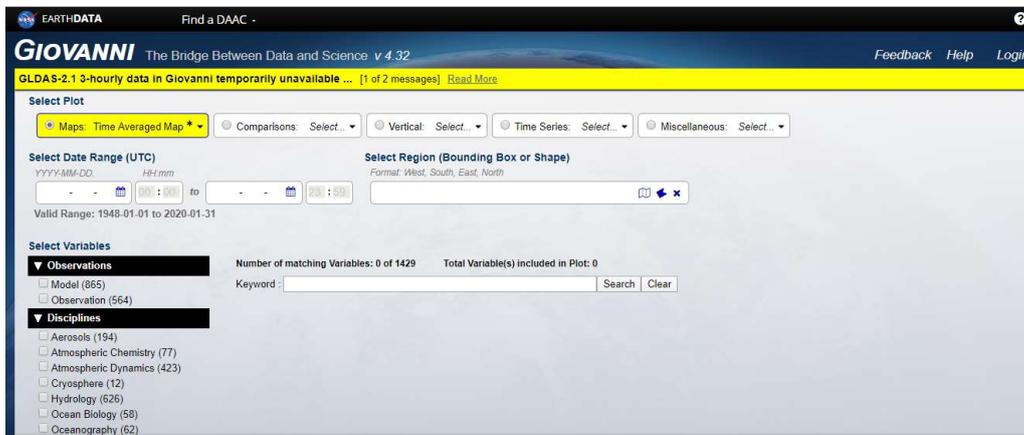


Figura 8 - Portal *Giovanni*. Fonte: <https://giovanni.gsfc.nasa.gov/giovanni/>

Muito dos conjuntos de dados existentes no Giovanni, principalmente os de alta resolução temporal (diariamente, a cada 3 ou 4 horas) podem apresentar falhas nos dados, uma descontinuidade no conjunto de dados durante um período. Esta falha pode dar-se por vários motivos, o motivo mais comum é que durante esse período o instrumento não estava a recolher dados, mas pode dar-se também devido a um problema na qualidade de dados, a uma quebra sazonal ou uma condição climática.

Os produtos disponíveis de interesse para este projeto, existentes no portal são mapas com o LAI e com a fração de verde na imagem, os dois produtos apresentam uma resolução espacial máxima de 75km. Estes produtos são gerados a partir de imagens obtidas pelo satélite MERRA-2 *Model* e o período de atualização é mensal. O acesso a estes produtos é gratuito.

Na **Figura 9** e na **Figura 10** são apresentados exemplos de *outputs* dos produtos mapas com o LAI e mapas com a fração de verde, respetivamente. Os produtos apresentados estão localizados em Portugal.

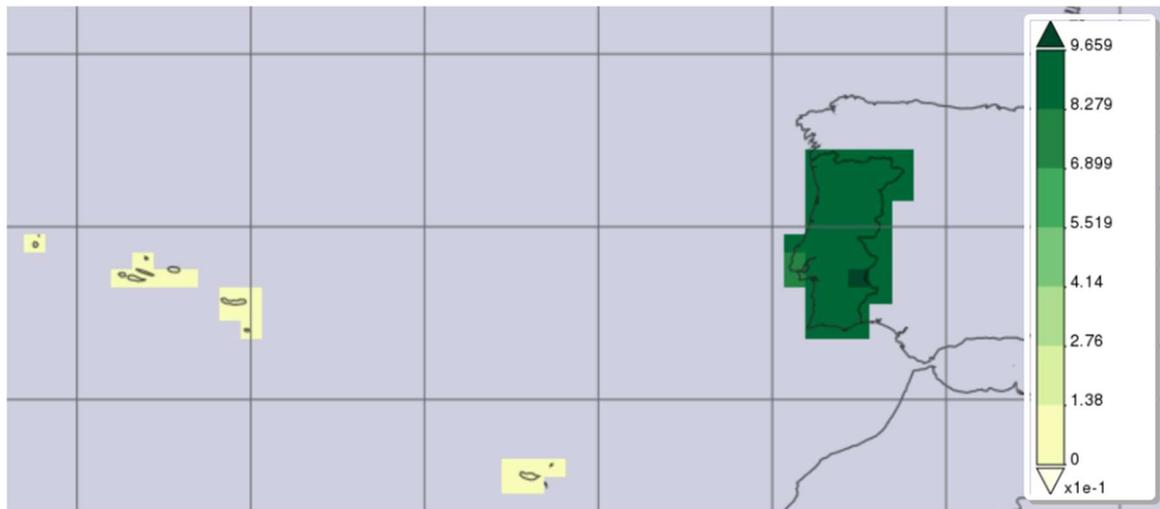


Figura 9 – Mapa com o LAI para Portugal.

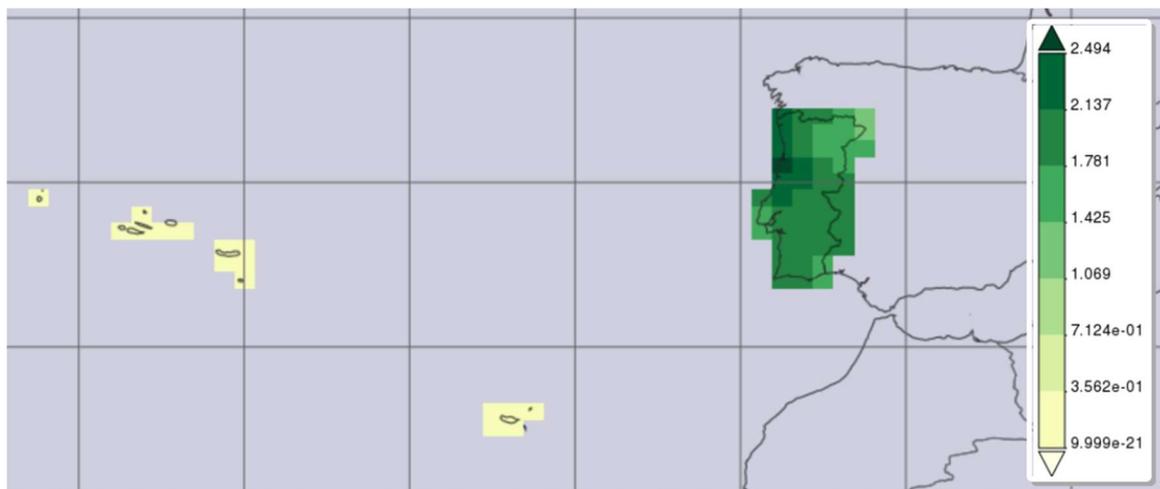


Figura 10 - Mapa com a versão de verde para Portugal.

3.4. *Land Cover*

O *Land Cover* (**Figura 11**) é um portal do *Copernicus Global Land Service* que possui mapas de cobertura do solo. Estes mapas podem ser dinâmicos, isto é, são incluídas transições de classes de cobertura do solo ao longo do tempo, permitindo assim conhecer-se as alterações na cobertura do solo. Para a obtenção dos produtos, o utilizador escolhe a área de interesse e seleciona o produto que deseja obter.

Os produtos globais *Land Cover* possuem um nível sobre o tipo de floresta para todos os pixéis em que a fração da cobertura florestal excede o 1%. Segundo o *Copernicus Global Land Service* (2020) há 6 tipos de classes para os produtos globais:

- Desconhecido/Outro,

- Floresta de folhas em agulha e verdes,
- Floresta de folhas largas e verdes,
- Floresta de folhas em agulha e caduca,
- Floresta de folhas largas e caduca,
- Misturada.



Figura 11 - Portal *Land Cover*. Fonte: <https://land.copernicus.eu/global/products/lc>

Os produtos disponíveis de interesse para este projeto, existentes no portal são mapas com os tipos de floresta, mapas de floresta aberta e fechada e mapas de floresta. Os produtos apresentam gratuitamente uma resolução espacial máxima de 10km. Estes produtos são gerados a partir de imagens obtidas pelo satélite *Sentinel 2* e apenas existem para o ano 2015. O acesso a estes produtos é gratuito.

Na **Figura 12** e na **Figura 13** são apresentados exemplos de *outputs* dos produtos: mapas com todos os tipos de floresta e mapas de floresta aberta e fechada, respetivamente. Os produtos apresentados foram selecionados para a zona de Abrantes, têm uma resolução espacial de 10km e são de 2015.

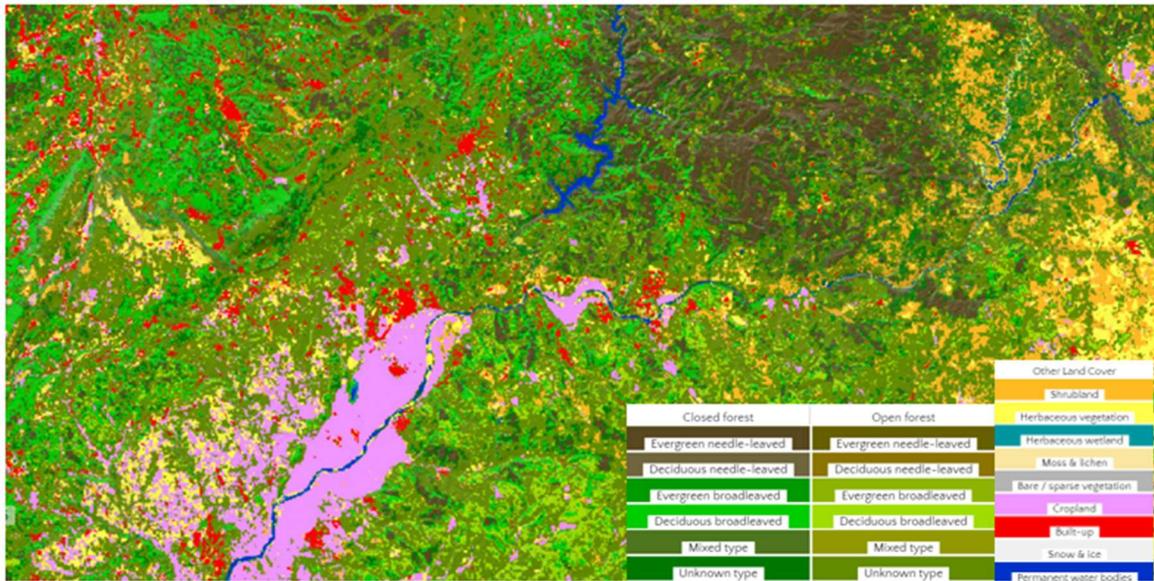


Figura 12 - Mapa com todos os tipos de floresta quando selecionada a zona Abrantes.

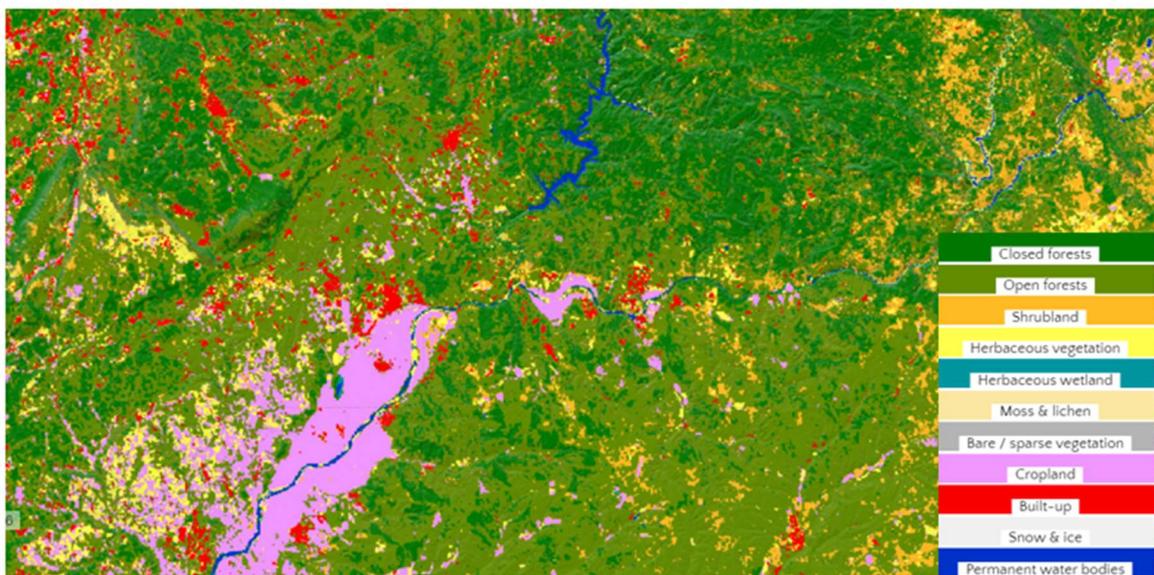


Figura 13 - Mapa com a floresta aberta e fechada quando selecionada a zona Abrantes.

3.5. *Global Forest*

O *Global Forest* (Figura 14) fornece dados e ferramentas para monitorizar florestas, de acesso gratuito. É possível escolher os produtos por datas, de modo a possibilitar comparações entre as mesmas. Tem como objetivo reunir os dados mais atuais, confiáveis e robustos para monitorizar as mudanças nas florestas em todo o mundo (*Global Forest Watch*, 2020).

A plataforma possui dados de:

- Alterações florestais (perda global de cobertura arbórea e dados de ganhos),
- Cobertura florestal,
- Incêndios florestais,
- Uso florestal,
- Dados de conservação (como limites globais para áreas protegidas e pontos críticos da biodiversidade),
- Dados de pessoas, incluindo os limites de terras comunitárias e os direitos de posse da terra (disponível em breve).



Figura 14 - Portal *Global Forest*. Fonte: <https://www.globalforestwatch.org/>

Os produtos disponíveis de interesse para este projeto, existentes no portal são mapas de cobertura arbórea e mapas de cobertura do solo, os dois produtos apresentam uma resolução espacial de 300m e 30m, respetivamente, quando é feito o *download*. Estes produtos são gerados a partir de imagens obtidas pelo satélite ESA e *Landsat*, respetivamente. Para o produto com a cobertura do solo só existem dados do ano 2015 e para o produto com a cobertura arbórea, o período de atualização é de 3 em 3 anos (entre o ano 2001 e o ano 2012). O acesso a estes produtos é gratuito.

Na **Figura 15** e na **Figura 16** são apresentados exemplos de *outputs* dos produtos: mapas de cobertura arbórea e mapas de cobertura do solo, respetivamente. Os produtos apresentados foram seleccionados para a zona de Abrantes. O mapa da cobertura arbórea (**Figura 15**) é do ano 2010 e o mapa da cobertura do solo (**Figura 16**) é de 2015.

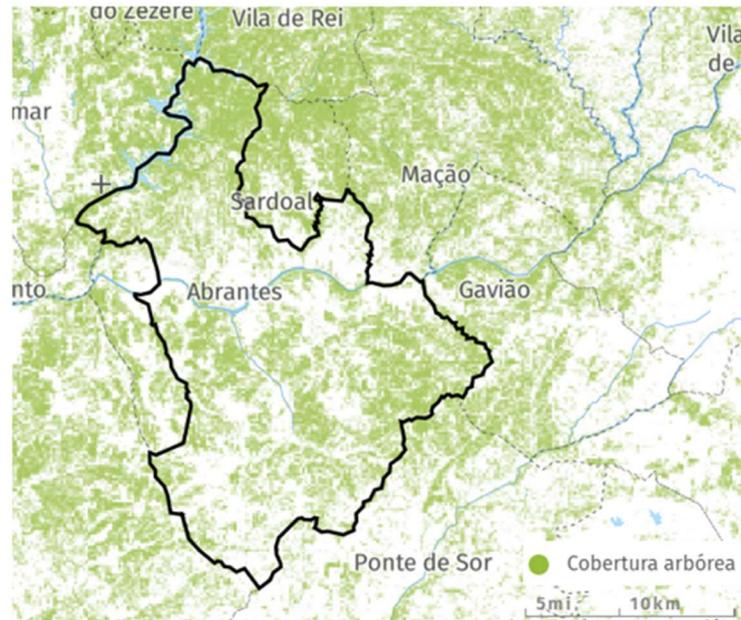


Figura 15 - Mapa da cobertura arbórea na zona de Abrantes.

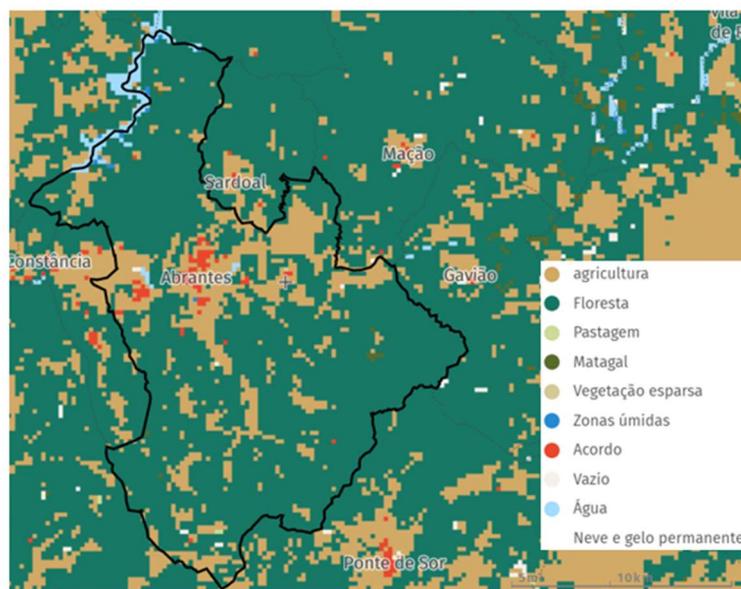


Figura 16 - Mapa da cobertura do solo na zona de Abrantes.

3.6. F-Tep

O F-Tep (Figura 17) é um portal que permite fácil acesso aos dados de satélite e às ferramentas para gerar produtos de informação florestal. Por meio da plataforma, os utilizadores “podem também criar e partilhar os seus próprios serviços, produtos e ferramentas geradas” (Forestry TEP, 2020). É uma plataforma aberta a utilizadores registados, autenticados por meio de uma conta ESA. Alguns dos benefícios que a plataforma pode oferecer são:

- Acesso a dados com funções fáceis de pesquisa,

- Acesso a produtos já disponíveis,
- Software pronto para uso sem necessidade de instalação,
- Serviços prontamente disponíveis,
- Balcão único para serviços de detecção remota florestal.

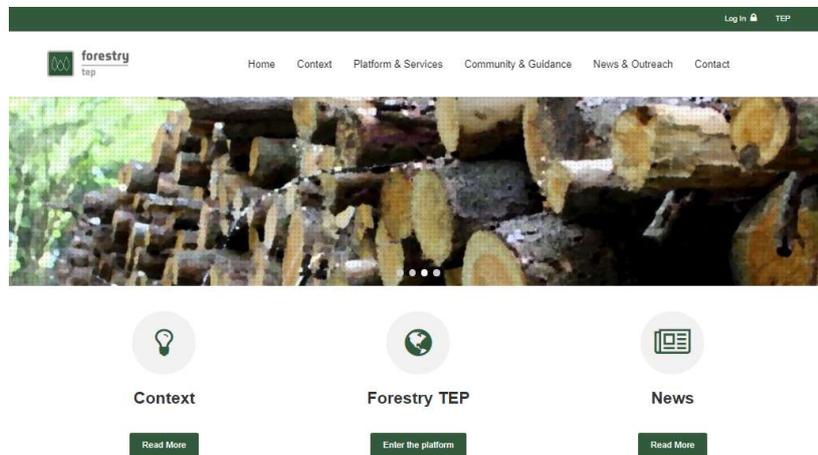


Figura 17 - Portal F-Tep. Fonte: <https://f-tep.com/>

A equipa de pesquisa *Technical Research Centre of Finland* (VTT) implementou no portal F-Tep serviços para a estimativa de parâmetros estruturais da floresta. Estes serviços utilizam dados do *Sentinel 2* e amostras de campo recolhidas pelo utilizador e posteriormente carregadas no portal, para estimar os parâmetros estruturais da floresta (como a contagem, o volume, o diâmetro, a área da base e altura de caules) de forma explícita com uma resolução de 10m.

O portal *Terramonitor* está a implementar serviços de processamento de dados no ambiente F-Tep. O enfoque principal das atividades do *Terramonitor* está no processamento de dados de satélites (*Sentinel 1*, *Sentinel 2* e *Sentinel 3*) no portal F-Tep, permitindo oferecer serviços globais. Futuramente, o processamento de dados poderá incluir a utilização dos dados do satélite *Landsat*, que também estão disponíveis na plataforma F-Tep. Esses processos fornecerão produtos de imagem sem nuvens, atualizados e fáceis de integrar em todo o mundo.

Esta plataforma integra as funcionalidades de um Sistema de Informação Geográfica (SIG), nomeadamente dos *softwares Quantum Geographic Information System (QGIS)*, *Monteverdi* e *Sentinel Application Platform (SNAP)*.

Os produtos disponíveis de interesse para este projeto, existentes no portal são mapas da alteração da floresta, mapas da cobertura do solo e mapas com índices de vegetação. Os produtos apresentam uma resolução espacial máxima de 10m. Estes produtos são gerados a partir de imagens obtidas pelo satélite *Sentinel 2* e o período de atualização destes é de 5 em 5 dias, se os dois satélites (S2A e S2B) trabalharem em simultâneo, ou de 10 em 10 dias se só for utilizado um deles. O acesso a estes produtos é gratuito, mas, limitado.

Exemplos de *outputs* dos produtos disponíveis de interesse são mostrados na secção 4.

3.7. Comparação das plataformas

Foram várias as plataformas exploradas e dessas foram selecionadas as que não seriam capazes de cumprir com os objetivos propostos. Razões para se ter rejeitado as seguintes plataformas:

- *Terramonitor*: Os produtos de interesse são serviços pagos, um dos objetivos era testar produtos de acesso gratuito que pudessem fornecer resultados qualitativos, só no caso de isso não ser possível é que se testariam serviços pagos;
- *Giovanni*: Os produtos de interesse apresentam baixa resolução e baixo período de atualização;
- *Sentinel Hub*: Os produtos de interesse apresentam baixa resolução e o acesso a produtos de alta resolução é pago;
- *Global Forest e Land Cover*: Os produtos de interesse apresentam baixa resolução e a última atualização dos produtos foi em 2015.

Tendo em conta as limitações referidas anteriormente foi escolhida a F-Tep para explorar detalhadamente. Com base na **Tabela 2**, que apresenta uma visão geral dos portais temos que, o portal F-Tep apresenta:

- Resolução espacial que possibilita efetuar classificações de mapas de ocupação do solo,
- Alto período de atualização,
- Produtos disponíveis de interesse para a classificação de imagens e para a identificação das alterações de floresta,
- Acesso gratuito, no entanto, limitado ao uso de 100 coins (máximo de moedas virtuais a usar gratuitamente).

Portais	Produtos disponíveis de interesse	Período de atualização	Resolução espacial	Preço	Classes	Fonte de dados	Obtenção dos produtos	Link para a aceder ao portal
Terramonitor	Vegetation (Vegetation heatmap of photosynthetic activity)	Mapas muito recentes (com um ou dois dias)	5m	-	Sem vegetação; vegetação escassa; vegetação média; vegetação densa	Sentinel 2	Camada WMS/WMTS	https://www.terramonitor.pt/
	VESA index	5 em 5 dias	Resolução standard: 10m x 10m Para mais detalhe disponibilizam (contactando-os): 5m, 3m, 1.5m e 50cm	Bronze: 199 €/mês Bronze (100ha-1000ha): 1€/ha; Silver (1000ha-5000ha): 0.85€/ha; Gold (5000ha-20000ha): 0.70€/ha; Platinum (>20000ha): é necessário contactá-los	Muito alto, alto, moderado, baixo e muito baixo Floresta, sem floresta			
Giovanni	Greenness fraction	Mensal	75km	Acesso gratuito	Stretched values	MERSA-2 Model	Download	https://giovanni.esf.edu.br/
	Leaf area index				Stretched values			
Forestry Tep	Forest Change	5 dias (os dois satélites forem em simultâneo) ou 10 dias (é usado um deles)	10m	Acesso gratuito	Classas definidas pelos dados de referência	Sentinel 2	Download	https://fisp.com/
	Land Cover				Stretched values			
Sentinel Hub	Vegetation Indices	5 dias (os dois satélites forem em simultâneo) ou 10 dias (é usado um deles)	100m	Uso individual não-comercial: 16.40€/mês; uso individual comercial: 99€/mês; desenvolvedores de app e empreendimento: desde 500€/mês	Stretched values	Sentinel 2	Download Camada WMS/WMTS	https://www.sentinel-hub.com/
	Agricultura				Não disponibilizam as classes, só os produtos sem legenda			
Global Forest	Índice de vegetação	Só tem do ano 2015	300m	Acesso gratuito	Agricultura, floresta, pastagem, matagal, vegetação esparsa, zonas húmidas, acórdo, vazio, água, neve e gelo permanentes	ESA Climate Change Initiative - Land Cover http://land.cci.earth.esa.int/	Download	https://www.globalforests.watch.org/
	Cobertura da solo				Ganho e perda da cobertura arbórea			
Land Cover	Cobertura arbórea	3 anos (2001-2012)	30m	Acesso gratuito		Landat	Download	https://land.copernicus.eu/
	All forest types							
	Open/closed forest	Só tem do ano 2015	10km	Acesso gratuito	- Para open/closed forest: evergreen needle-leaved, deciduous needle-leaved, evergreen closed forests, open forests, shrubland, herbaceous vegetation, herbaceous wetland, moss & lichen, bare/sparse vegetation, cropland, built-up, snow & ice, permanent water bodies	Sentinel 2	Download	https://land.copernicus.eu/
	Single class forest				Forests, shrubland, herbaceous vegetation, herbaceous wetland, moss & lichen, bare/sparse vegetation, cropland, built-up, snow & ice, permanent water bodies			

Tabela 2 - Quadro geral das plataformas estudadas.

4. Portal F-Tep

Neste capítulo é apresentado com maior detalhe o portal F-Tep, são listados todos os produtos e ferramentas disponíveis bem como o seu funcionamento. Em virtude das limitações de tempo para a realização deste projeto decidiu-se explorar com maior detalhe apenas o processamento das imagens do *Sentinel 2*.

4.1. Produtos disponíveis

Os produtos disponíveis na plataforma F-Tep são apresentados nesta secção, tanto para produtos que utilizam o processamento de imagens do *Sentinel 2* (secção 4.1.1), de imagens do *Sentinel 1* (secção 4.1.2) e de imagens de outros satélites (secção 4.1.3).

Para além dos produtos detalhados nas secções 4.1.1., 4.1.2. e 4.1.3, a plataforma apresenta também serviços de aplicação como o Monteverdi, o QGIS v2.18, o QGIS v3.6 e o SNAP. Estes serviços permitem a navegação em cada uma das aplicações, descritas anteriormente, num separador *web* durante o tempo definido pelo utilizador (sendo o máximo 1 hora). Estes serviços são úteis para quem precisa de trabalhar nestes esporadicamente e durante pouco tempo, não sendo necessária a instalação das aplicações indicadas anteriormente, tendo em conta que podem ser utilizadas através do *browser*.

Para todos estes serviços de aplicação o único dado necessário é uma imagem ou o produto que o utilizador deseja abrir na aplicação. Os parâmetros de entrada são os mesmos em todas as aplicações: a imagem de satélite que se pretende utilizar, o tempo limite de utilização antes da aplicação desligar automaticamente e o nome do trabalho.

4.1.1. Para imagens do Sentinel-2

Nesta secção serão apresentados os serviços disponíveis na plataforma F-Tep para o processamento de imagens do *Sentinel 2*, será apresentado para cada serviço: a sua definição, os dados e os parâmetros necessários para correr o serviço e, por fim são apresentados exemplos de *outputs*.

Em quase todas as ferramentas para o processamento de imagens do *Sentinel 2* o dado de entrada é sempre uma imagem do Nível-1C. Imagens deste nível não apresentam correções relativas à atmosfera, no entanto, o algoritmo presente em cada ferramenta do portal faz essas correções automaticamente, exceto na ferramenta “*VegetationIndices*” como pode ser visto na secção 4.1.1.2.

4.1.1.1. Serviço LandCoverS2

O serviço “*LandCoverS2*” permite o mapeamento da cobertura do solo utilizando o algoritmo *Random Forest*, que se encontra implementado na *Orfeo Toolbox* (OTB) (Miettinen, 2019) e que é explicado no capítulo 2, na secção 2.2.

Para este serviço os dados de entrada são: uma imagem do *Sentinel 2* Nível-1C e uma *shapefile*, que contém os dados de referência para o treino do classificador. Como parâmetros necessários para correr o serviço temos:

- Imagem de satélite que se pretende utilizar,
- Identificador uniforme de recurso (IUR) da pasta zipada com os dados de treino na secção “*Files*”,
- Nome do campo presente na *shapefile* que tem a classificação das amostras,
- Especificação do Sistema de Coordenadas de Referência (SCR) da área de interesse,
- Área de interesse (esta é seleccionada ou inserindo as coordenadas de cada ponto ou por um polígono que delimita a área desejada no sistema de coordenadas seleccionado no parâmetro anterior),
- Resolução espacial pretendida (é escolhida a resolução do produto que será gerado, podendo variar entre 10 a 100m),
- Nome do trabalho.

Como *output* temos: uma imagem com o mapeamento da cobertura do solo com as classes definidas nos dados de treino em formato *.tif, a matriz de confusão em formato *.csv e o modelo de treino em formato *.txt. Todos podem ser obtidos pelo custo de 1 *coin*.

Um exemplo de *output* ao correr este serviço é mostrado na **Figura 18**.

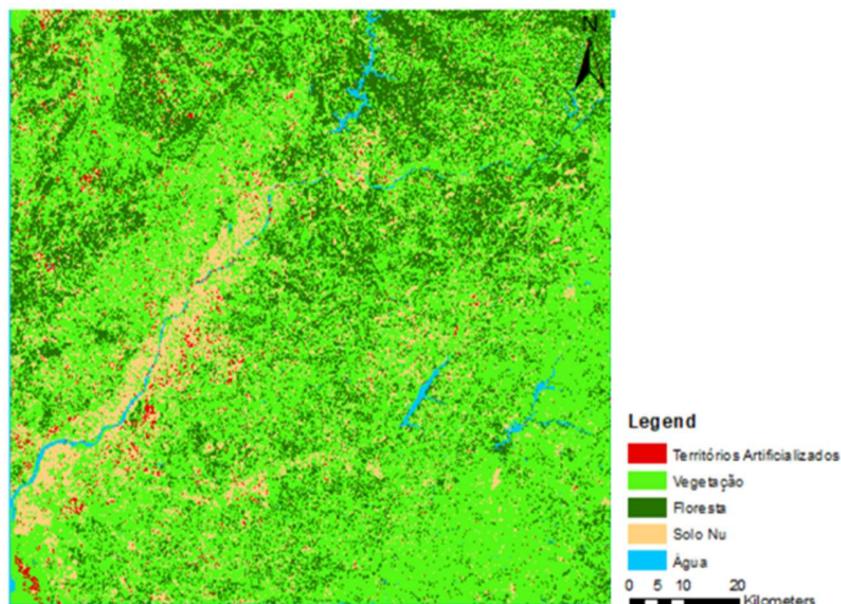


Figura 18 - Exemplo de *output* da ferramenta LandCoverS2

4.1.1.2. Serviço *VegetationIndices*

O serviço “*VegetationIndices*” calcula o índice de vegetação de uma determinada área de interesse para imagens obtidas pelo satélite *Sentinel 2* (Miettinen, 2019).

Para este serviço, o único dado necessário é uma imagem do *Sentinel 2*. Devem ser escolhidas imagens sem nuvens, pois o serviço não dispõe dessas correções. Este serviço calcula o índice de vegetação com o NDVI. O conceito NDVI é explicado no capítulo 2, na secção 2.3. Como parâmetros necessários para correr o serviço temos:

- Imagem de satélite que se pretende utilizar,
- Especificação do SCR da área de interesse,
- Área de interesse (esta é seleccionada ou inserindo as coordenadas de cada ponto ou por um polígono que delimita a área desejada, no sistema de coordenadas seleccionado no parâmetro anterior),
- Resolução espacial (é escolhida a resolução do produto que será gerado, podendo variar entre 10 a 100m),
- Nome do trabalho.

Como *output* temos: uma imagem em formato *.tif com o índice de vegetação da área de interesse escolhida.

Um exemplo de *output* ao correr este serviço é mostrado na **Figura 19**. Os valores são apresentados numa escala gradual de cores, em que a preto o índice de vegetação é baixo e a branco o índice de vegetação é alto.

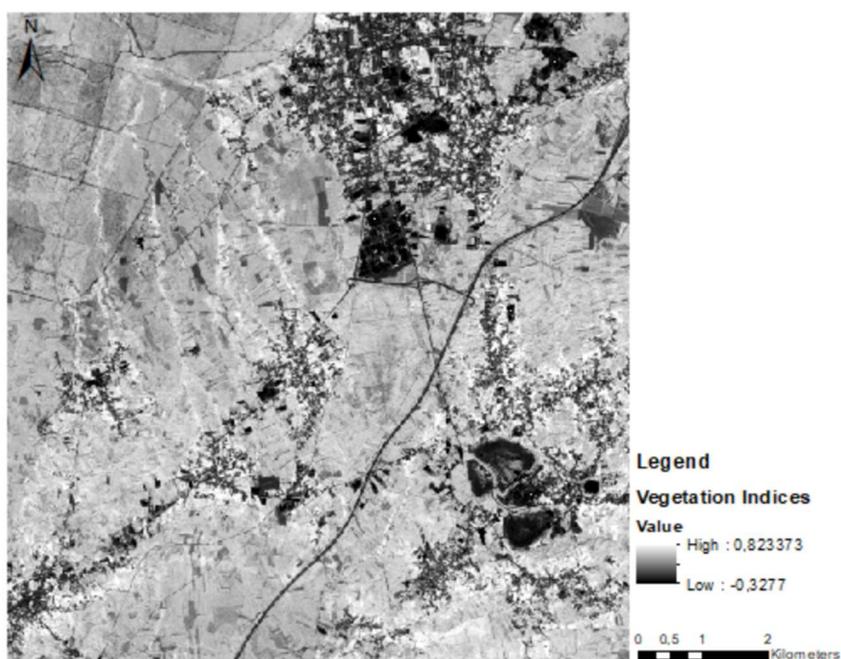


Figura 19 - Exemplo de *output* da ferramenta *VegetationIndices*

4.1.1.3. Serviço *VegetationIndicesOLD*

O serviço “*VegetationIndicesOLD*” calcula o índice de vegetação de uma determinada área de interesse. Este serviço utiliza o antigo formato de convenção dos nomes dos produtos gerados pelas imagens obtidas pelo satélite *Sentinel 2* (Miettinen, 2019), este formato é abordado no capítulo 2, na secção 2.1.

Para este serviço os dados de entrada são: uma imagem do *Sentinel 2* Nível-1C. Este serviço pode calcular o índice de vegetação, segundo a preferência do utilizador, utilizando os seguintes índices: NDVI, TNDVI, RVI, SAVI, TSAVI, MSAVI, MSAVI2, GEMI e IPVI. Como parâmetros necessários para correr o serviço temos:

- Imagem de satélite que se pretende utilizar,
- Escolha do índice de interesse,
- Especificação do SCR da área de interesse,
- Área de interesse (esta é selecionada ou inserindo as coordenadas de cada ponto ou por um polígono que delimita a área desejada, no sistema de coordenadas selecionado no parâmetro anterior),
- Resolução espacial (é escolhida a resolução do produto que será gerado, podendo variar entre 10 a 100m),
- Nome do trabalho.

Como *output* temos: uma imagem em formato *.tif com o índice de vegetação da área de interesse escolhida nos parâmetros, os valores são apresentados numa escala gradual de cores, em que a preto o índice de vegetação é baixo e a branco o índice de vegetação é alto.

Um exemplo de *output* neste serviço é como o que é mostrado na **Figura 19**, só altera o índice com que é calculado o índice de vegetação.

4.1.1.4. Serviço *ForestChangeS2*

O serviço “*ForestChangeS2*” permite observar as mudanças ocorridas na floresta de interesse. Trabalha com a diferença da refletância dos materiais na banda do vermelho (Miettinen, 2019).

Para este serviço os dados de entrada são: duas imagens do *Sentinel 2* Nível-1C. As duas imagens devem ocupar a mesma área. Como parâmetros necessários para correr o serviço temos:

- Primeira imagem adquirida antes ou durante a alteração da floresta,
- Última imagem adquirida depois ou durante a alteração da floresta,
- Especificação do SCR da área de interesse,

- Área de interesse (esta é selecionada ou inserindo as coordenadas de cada ponto ou por um polígono que delimita a área desejada, no sistema de coordenadas selecionado no parâmetro anterior),
- Resolução espacial (é escolhida a resolução do produto que será gerado, podendo variar entre 10 a 100m),
- Nome do trabalho.

Como *output* temos: uma imagem em formato *.tif com a alteração da floresta da área de interesse escolhida nos parâmetros. Os valores são apresentados numa escala gradual de cores, em que a preto a alteração é baixa e a branco a alteração é alta.

Um exemplo de *output* ao correr este serviço é mostrado na **Figura 20**.

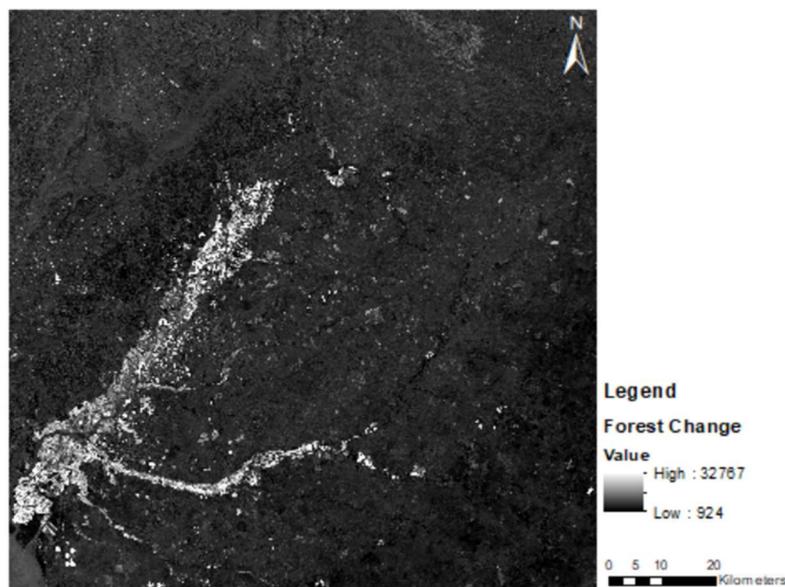


Figura 20 - Exemplo de *output* da ferramenta ForestChangeS2

4.1.1.5. Serviço CombS2granules & CombS2granulesZ

Segundo Miettinen (2019), o formato *Sentinel 2* padrão divide cada imagem em blocos (mais conhecido por *tiles*) de 110 km por 110 km. Estes serviços permitem combinar esses blocos num único ficheiro *geotiff*, para análise posterior como uma única unidade. A forma básica “*CombS2granules*” só aceita blocos na mesma zona da projeção *Universal Transversal de Mercator* (UTM) e a versão “*CombS2granulesZ*” aceita blocos de várias zonas UTM, reprojetoando-os no sistema de coordenadas de saída definido pelo utilizador.

Para estes serviços os dados de entrada são: pelo menos dois ou mais blocos do *Sentinel 2* Nível-1C. Como parâmetros necessários para correr o serviço temos:

- Imagem de satélite que se pretende utilizar,

- Especificação do SCR da área de interesse,
- Área de interesse (esta é seleccionada ou inserindo as coordenadas de cada ponto ou por um polígono que delimita a área desejada, no sistema de coordenadas seleccionado no parâmetro anterior),
- Início do nome do ficheiro *geotiff* de saída,
- Resolução espacial (é escolhida a resolução do produto que será gerado, podendo variar entre 20 a 100m),
- Nome do trabalho.

Como *output* temos uma imagem em formato *.tif com os blocos juntos num só ficheiro.

Na **Figura 21** é mostrado um *output* da ferramenta CombS2granules, que apresenta a junção de dois blocos na mesma zona UTM. Na **Figura 22** é mostrado um *output* da ferramenta CombS2granulesZ, que apresenta a junção de dois blocos de diferente zonas UTM.

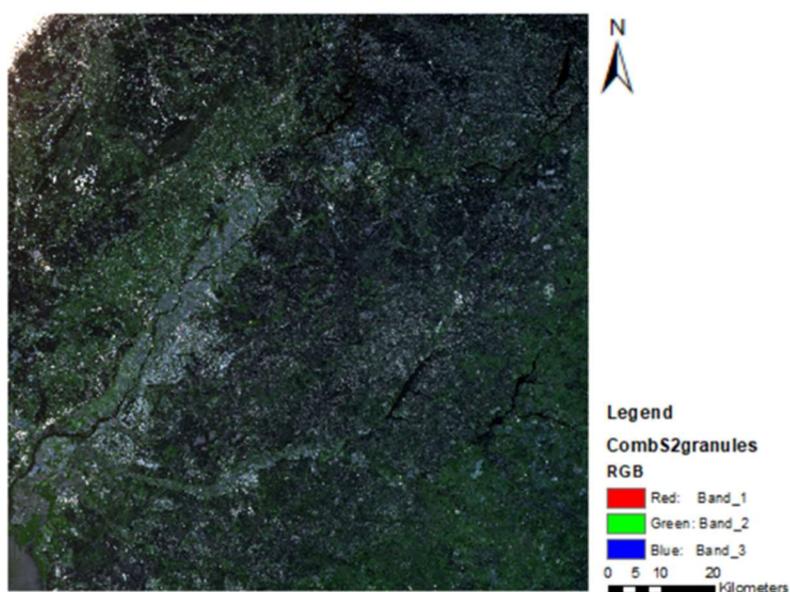


Figura 21 - Exemplo de *output* da ferramenta CombS2granules

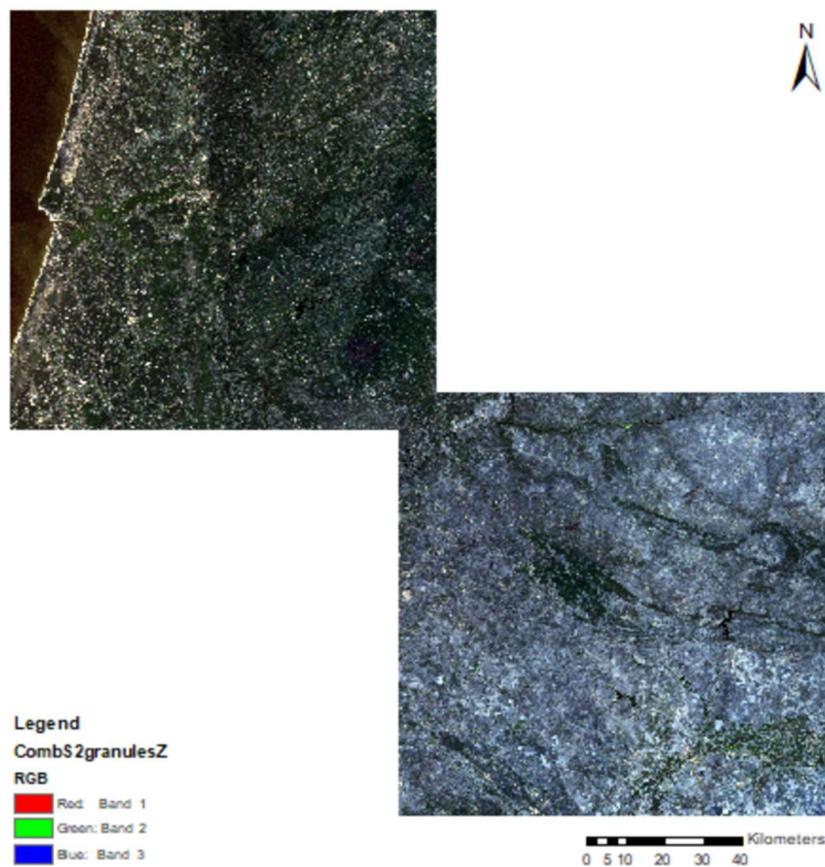


Figura 22 - Exemplo de output da ferramenta CombS2granulesZ

4.1.2. Para imagens do Sentinel-1

Nesta secção serão apresentados os serviços disponíveis na plataforma F-Tep para o processamento de imagens do *Sentinel 1*. Como este satélite não foi o enfoque do trabalho será dada apenas uma definição, os dados e parâmetros necessários para correr os serviços, não serão apresentados exemplos de *outputs* em cada ferramenta.

4.1.2.1. Serviço LandCoverS1

Segundo Miettinen (2019), o serviço “*LandCoverS1*” permite o mapeamento da cobertura do solo com séries temporais do *Sentinel 1*. As séries temporais permitem a análise ao longo do tempo de determinada área, a combinação das várias imagens permite o mapeamento dos tipos de cobertura do solo, utilizando o algoritmo *Random Forest*, que se encontra implementado na OTB e é explicado no capítulo 2, na secção 2.2.

Para este serviço os dados de entrada são: uma imagem do *Sentinel 1* e uma *shapefile* com os dados de referência. Como parâmetros necessários para correr o serviço temos:

- Imagem de satélite que se pretende utilizar,
- IUR da pasta zipada com os dados de treino na secção “*Files*”,
- Nome do campo presente na *shapefile* que tem a classificação das amostras,
- Especificação do SCR da área de interesse,
- Área de interesse (esta é seleccionada ou inserindo as coordenadas de cada ponto ou por um polígono que delimita a área desejada, no sistema de coordenadas seleccionado no parâmetro anterior),
- DEM a utilizar,
- Resolução espacial (é escolhida a resolução do produto que será gerado, podendo variar entre 10 a 100m),
- Nome do trabalho.

Como *output* temos: uma imagem com o mapeamento da cobertura do solo com as classes definidas nos dados de treino em formato *.tif, a matriz de confusão em formato *.csv e o modelo de treino em formato *.txt.

4.1.2.2. Serviço S1Biomass

O serviço “*S1Biomass*” permite o mapeamento da biomassa existente numa determinada área, quanto maior a quantidade de biomassa de uma área florestal maior será a refletância dos materiais (Miettinen, 2019).

Para este serviço os dados de entrada são: um DEM, uma máscara de floresta e uma coleção de imagens do *Sentinel 1*. Como parâmetros necessários para correr o serviço temos:

- Imagem de satélite que se pretende utilizar,
- DEM de referência,
- Máscara de floresta de referência,
- Especificação do SCR da área de interesse,
- Área de interesse (esta é seleccionada ou inserindo as coordenadas de cada ponto ou por um polígono que delimita a área desejada, no sistema de coordenadas seleccionado no parâmetro anterior),
- Resolução espacial (é escolhida a resolução do produto que será gerado, podendo variar entre 10 a 100m),
- Nome do trabalho.

Como *output* temos uma imagem em formato *.tif em que os valores são apresentados numa escala gradual de cores, valores altos indicam alta quantidade de biomassa e valores baixos baixa quantidade de biomassa.

4.1.2.3. Serviço S1stack

Segundo Miettinen (2019), o serviço “*S1stack*” faz mosaicos de imagens *Ground Range Detected* (GRD) do *Sentinel 1* com polarização dupla ou bipolar (A polarização refere-se à direção de propagação das ondas eletromagnéticas, esta pode ser polarizada verticalmente e/ou horizontalmente, no caso da polarização dupla, o sensor recebe e transmite os dois tipos de polarização) ao longo da órbita e atribui datas de aquisição como bandas em uma imagem de saída multitemporal.

Para este serviço os dados de entrada são: um conjunto de imagens *Sentinel 1* bipolares com mesma geometria no formato GRD como dados de entrada. Como parâmetros necessários para correr o serviço temos:

- Conjunto de imagens *Sentinel 1*,
- Sistema de coordenadas,
- Área de interesse (esta é selecionada ou inserindo as coordenadas de cada ponto ou por um polígono que delimita a área desejada, no sistema de coordenadas selecionado no parâmetro anterior),
- DEM de referência,
- Nome do *output*,
- Resolução espacial (é escolhida a resolução do produto que será gerado, podendo variar entre 20 a 100m),
- Nome do trabalho.

Como *output* temos uma imagem em formato *.tif com o conjunto de mosaicos de imagens em formato GRD.

4.1.3. Para imagens de qualquer satélite

Nesta secção é apresentado o serviço disponível na plataforma F-Tep para o processamento de imagens de qualquer satélite disponível no portal F-Tep.

4.1.3.1. Serviço LandCoverGeotiff

O serviço “*LandCoverGeotiff*” faz a classificação da cobertura do solo utilizando o algoritmo *Random Forest* (Miettinen, 2019), explicado no capítulo 2, na secção 2.2.

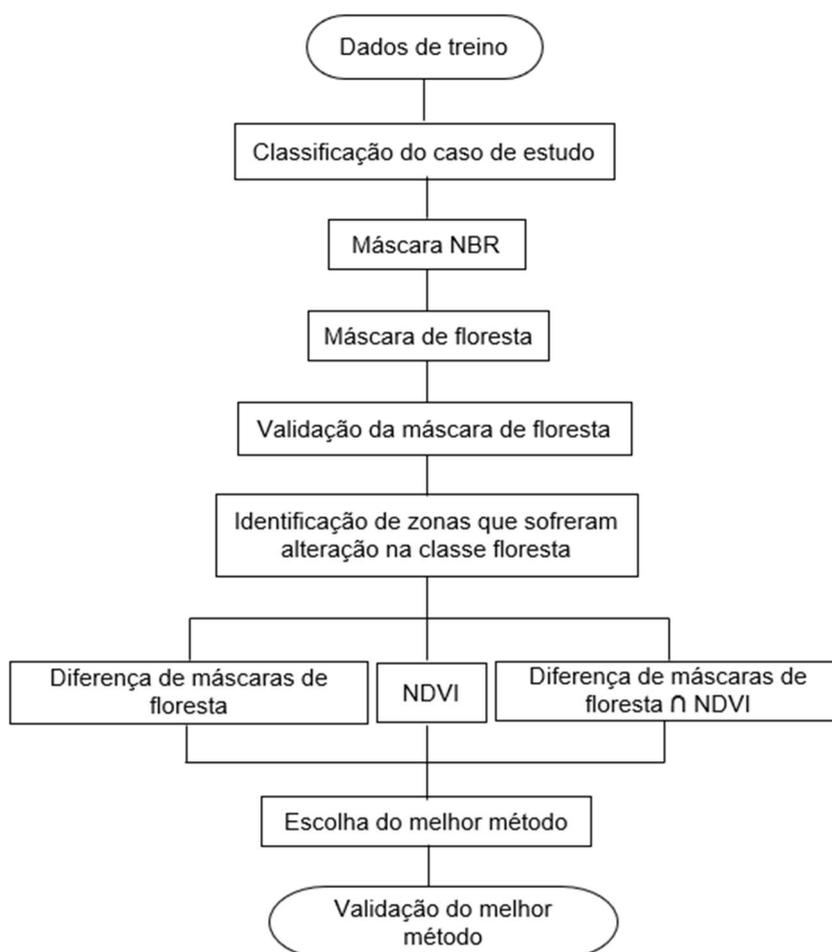
Para este serviço o único dado necessário é uma imagem *geotiff* como ficheiro de entrada, que pode ser de qualquer sensor de satélite ou uma combinação de vários sensores. Como parâmetros necessários para correr o serviço temos:

- Imagem de satélite que se pretende utilizar,
- IUR da pasta zipada com os dados de treino na secção “*Files*”,
- Nome do campo presente na *shapefile* que tem a classificação das amostras,
- Especificação do SCR da área de interesse,
- Área de interesse (esta é seleccionada ou inserindo as coordenadas de cada ponto ou por um polígono que delimita a área desejada, no sistema de coordenadas seleccionado no parâmetro anterior),
- Resolução espacial (é escolhida a resolução do produto que será gerado, podendo variar entre 10 a 100m),
- Nome do trabalho.

Como *output* temos: uma imagem com o mapeamento da cobertura do solo com as classes definidas nos dados de treino em formato *.tif, a matriz de confusão em formato *.csv e o modelo de treino em formato *.txt.

5. Identificação automática das zonas de floresta e de alteração de floresta

No capítulo 5 faz-se a descrição do trabalho prático desenvolvido no âmbito desta tese, que tem enfoque na identificação automática de florestas e de alterações de floresta, presentes no caso de estudo apresentado na secção 5.1. O **Fluxograma 1** mostra os passos gerais para o cumprimento dos objetivos referidos anteriormente.



Fluxograma 1 - Síntese metodológica dos passos desenvolvidos na parte prática

5.1. Caso de Estudo

O RAIZ sugeriu quatro potenciais áreas de estudo (**Figura 23**) localizadas no continente português. Seguindo a Nomenclatura de Unidades Territoriais de Estatística (NUTS), estas áreas estão localizadas no NUTS 2.

A zona 1, designada por “Quintarrei” (RAIZ1), está localizada na região norte de Portugal. As zonas 2 e 3, designadas por “Baixo Vouga” (RAIZ2) e “Caniceira” (RAIZ3) respetivamente, estão localizadas na região centro de Portugal. A zona 4, designada por “Área de Intervenção Florestal (ZIF) da Calha do Grou” (RAIZ4), está localizada na região do Alentejo.

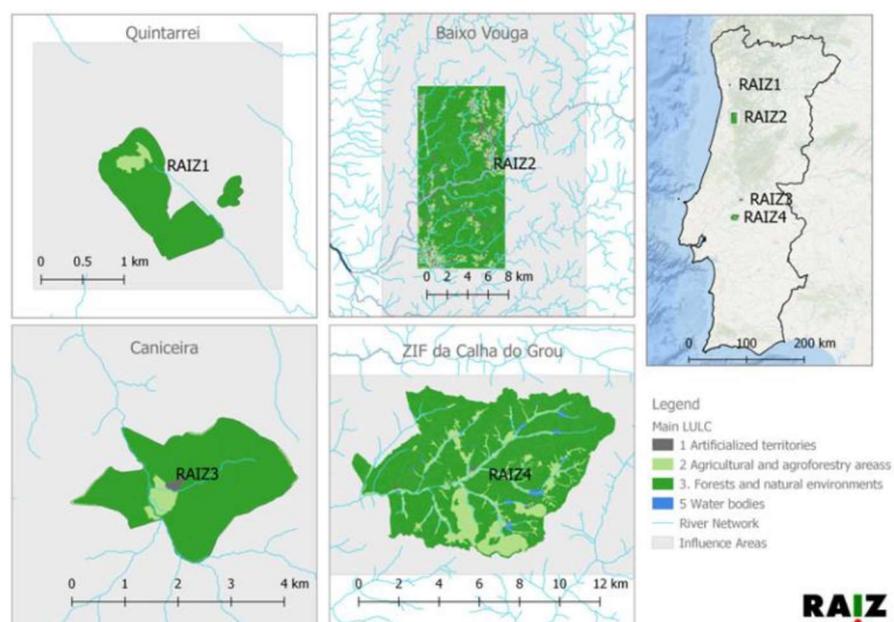


Figura 23 – Áreas de interesse no território português. Fonte: <https://mysustainableforest.com/case-demos/portugal/>

Das quatro zonas apresentadas, a zona 3 foi a escolhida para a realização deste projeto. A RAIZ3 localiza-se na zona da Caniceira, uma aldeia pertencente à freguesia de Vale de Cavalos e ao concelho da Chamusca. Esta zona é gerida pela *The Navigator Company* e tem como espécie dominante o *Eucalyptus globulus*, da família do eucalipto. Relativamente ao clima, é uma zona que se encontra numa região com verão seco e quente.

Foi feita uma análise sobre a área de interesse proposta, no entanto, e tendo em conta o trabalho que seria feito sobre a mesma, achou-se que seria muito mais vantajoso estender a área de estudo para a imagem de satélite em que a zona RAIZ3 está inserida, como pode ser visto na **Figura 24**.



Figura 24 - Caso de Estudo.

5.2. Dados de referência

Os dados de referência utilizados para a validação da máscara de floresta e da máscara de alterações de floresta são dados fornecidos pelo RAIZ, recolhidos em campo por *Global Positioning System* (GPS). Dos dados fornecidos apenas foram utilizados os dados:

- De um inventário florestal realizado na zona RAIZ3, recolhidos a 12 de fevereiro de 2020;
- De validação da máscara de floresta realizado na zona RAIZ3, recolhidos em março de 2019;

Os dados relativos ao inventário florestal são compostos por 122 pontos, nestes pontos está a informação do tipo de floresta, dos tipos de cortes na floresta (por exemplo: corte raso de pinheiro), dos matos, dos territórios artificializados, da agricultura e da água.

Os dados de validação da máscara de floresta são compostos por 538 pontos, com a informação do tipo de floresta e de sítios sem floresta.

Idealmente deveria ter sido construída uma base de dados de referência usando um processo de amostragem como os que são explicados na secção 2.4. mas em virtude das limitações de tempo para a realização do projeto optou-se por usar apenas os dados recolhidos pelo RAIZ.

5.3. Dados de treino

Os dados de treino, com geometria do tipo polígono, foram criados no ArcGIS. As imagens utilizadas para a criação dos dados de treino e posterior classificação foram recolhidas no verão, a 3 de agosto de 2019, no outono, a 27 de setembro de 2019 e no inverno, a 10 de março de 2020.

Para gerar os dados de treino foram identificados vários polígonos representativos de cada classe, através de observação visual da imagem a classificar e de imagens de muito alta resolução (*basemap* do ArcGIS). Para a criação destes foi preciso *à priori* a definição da nomenclatura a considerar para a ocupação do solo. Foi considerada a seguinte nomenclatura:

- Classe 1: Territórios artificializados,
- Classe 2: Vegetação herbácea e arbustiva,
- Classe 3: Floresta,
- Classe 4: Solo nu,
- Classe 5: Água.

A nomenclatura escolhida foi inspirada na correspondência entre a nomenclatura *Corine Land Cover* (CLC) e as megaclasses temáticas, que pode ser vista na **Figura 25**. Como o enfoque era na floresta algumas das megaclasses foram alteradas e outras não foram consideradas.

Nomenclatura CLC	Megaclases
111 Tecido urbano contínuo	Territórios artificializados
112 Tecido urbano descontínuo	
121 Indústria, comércio e equipamentos gerais	
122 Redes viárias e ferroviárias e espaços associados	
123 Áreas portuárias	
124 Aeroportos e aeródromos	
131 Áreas de extracção de inertes	
132 Áreas de deposição de resíduos	
133 Áreas em construção	
141 Espaços verdes urbanos	
142 Equipamentos desportivos, culturais e de lazer e zonas históricas	
211 Culturas temporárias de sequeiro	Agricultura
212 Culturas temporárias de regadio	
213 Arrozais	
221 Vinhas	
222 Pomares	
223 Olivais	
231 Pastagens permanentes	
241 Culturas temporárias e/ou pastagens associadas a culturas permanentes	Agricultura com áreas naturais e sistemas agro-florestais
242 Sistemas culturais e parcelares complexos	
243 Agricultura com espaços naturais e semi-naturais	Floresta
244 Sistemas agro-florestais	
311 Florestas de folhosas	Vegetação natural
312 Florestas de resinosas	
313 Florestas mistas	Floresta
321 Vegetação herbácea natural	
322 Matos	Espaços descobertos ou com vegetação esparsa
323 Vegetação esclerófila	
324 Florestas abertas, cortes e novas plantações	Floresta
331 Praias, dunas e areais	
332 Rocha nua	Floresta
333 Vegetação esparsa	
334 Áreas ardidas	Zonas húmidas
411 Paúis	
412 Turfeiras	
421 Sapais	Corpos de água
422 Salinas e aquicultura litoral	
423 Zonas entre-marés	Corpos de água
511 Cursos de água	
512 Planos de água	
521 Lagoas costeiras	
522 Desembocaduras fluviais	
523 Oceano	

Figura 25 - Correspondência entre a nomenclatura CLC e as megaclases. Fonte: <http://mapas.dgterritorio.pt/atom-dgt/pdf-cous/CLC2012/CLC-1990-2000-2006-2012-PT.pdf>

Os dados de treino foram criados para toda a imagem de satélite a que a zona de Abrantes pertence. Inicialmente, foram criados os dados de treino para a imagem de inverno (no ambiente ArcGIS), para os restantes períodos foram utilizados os mesmos dados só que eram eliminados os polígonos que não correspondiam à verdade nesse determinado período. Assim o utilizador não necessitaria de criar de raiz os dados de treino, eliminaria apenas os polígonos que para a sua imagem não fizessem sentido. Na **Figura 26** são mostrados exemplos de dados de treino criados para a imagem de inverno.

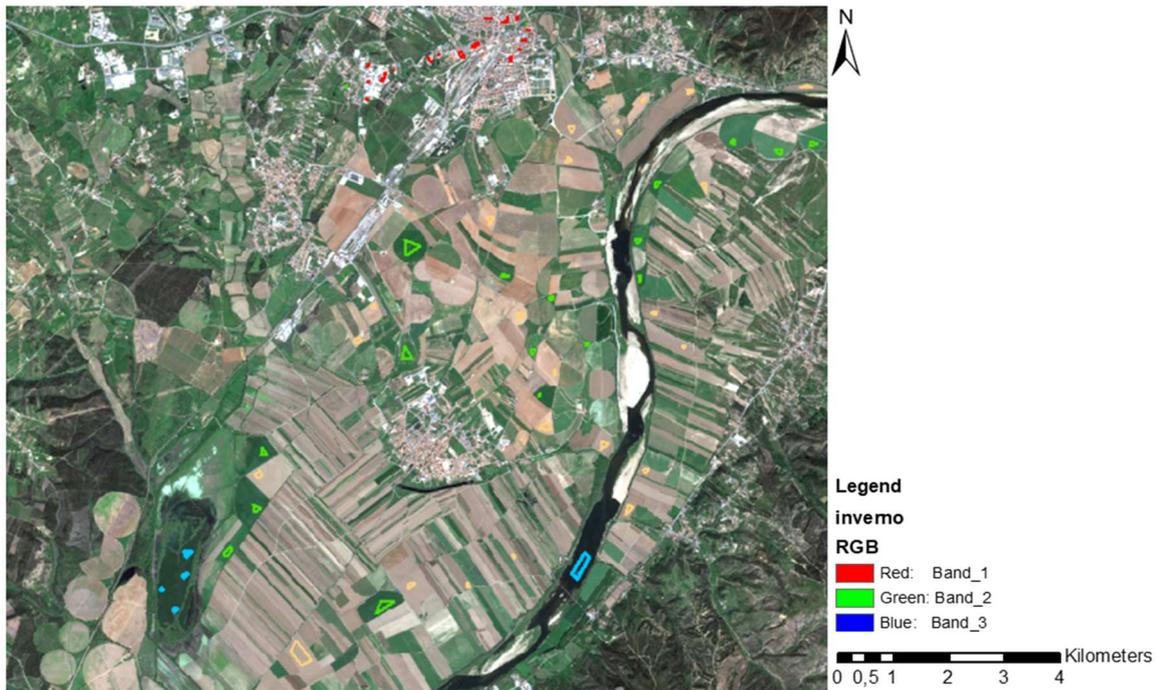


Figura 26 - Dados de treino da imagem de inverno.

Foi feita a análise da separabilidade das classes, que está relacionado com o quanto a resposta espectral das classes se encontram separadas umas das outras nos dados de treino. Caso haja uma má separabilidade entre as classes, estas vão ser confundidas na classificação. Foram gerados gráficos de dispersão para poder observar quais as classes que estariam a ser mais confundidas. Quanto maior a sobreposição de classes menor é a sua separabilidade. Um gráfico de dispersão permite analisar a relação entre duas variáveis quantitativas, uma variável y que resulta da aplicação de uma função a outra variável x .

Neste caso, mostram-se resultados para 3 variáveis: rgbc1 (banda 1 = cor vermelha), rgbc2 (banda 2 = cor verde) e rgbc3 (banda 3 = cor azul). Para poder mostrar a relação entre estas 3 variáveis as combinações possíveis são:

- rgbc1 e rgbc2,
- rgbc1 e rgbc3,
- rgbc2 e rgbc3.

Estas combinações irão representar os eixos x e y nos gráficos de dispersão apresentados na **Figura 27**, na **Figura 28** e na **Figura 29**. Para a compreensão do que é dito em seguida, é preciso ter em conta que:

- Classe 1 – Territórios Artificializados
- Classe 2 – Vegetação herbácea e arbustiva
- Classe 3 – Floresta

- Classe 4 – Solo Nu
- Classe 5 – Água

A **Figura 27** (a) apresenta uma visão geral da separabilidade das classes para a imagem de inverno consoante as variáveis consideradas. Para a classificação feita com uma imagem recolhida neste período, as classes que se mostraram com menor separabilidade foram as classes 1 e 4, as classes 2 e 5 e as classes 3 e 5.

A separabilidade das classes foi avaliada, e o processo de obtenção das áreas a usar foi feito por iterações, retirando ou adicionando polígonos às classes de forma a obter separabilidade suficiente para fazer a classificação. A **Figura 27** (b) mostra a separabilidade das classes depois da alteração dos dados de treino.

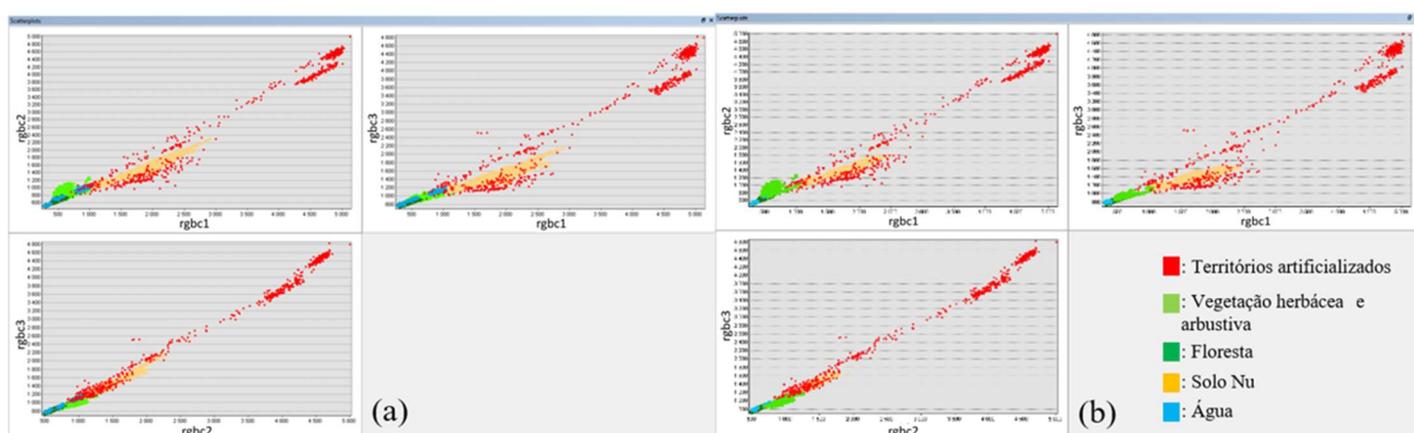


Figura 27 – Separabilidade das classes para a imagem de inverno. (a) Antes da alteração dos polígonos (b) Depois da alteração dos polígonos

A **Figura 28** (a) apresenta uma visão geral da separabilidade das classes para a imagem de outono consoante as variáveis consideradas. Para a classificação feita com uma imagem recolhida neste período, as classes que se mostraram com menor separabilidade foram as classes 1 e 2, as classes 1 e 4, as classes 2 e 4, as classes 2 e 5 e as classes 3 e 5. A **Figura 28** (b) mostra a separabilidade das classes depois da alteração dos dados de treino.

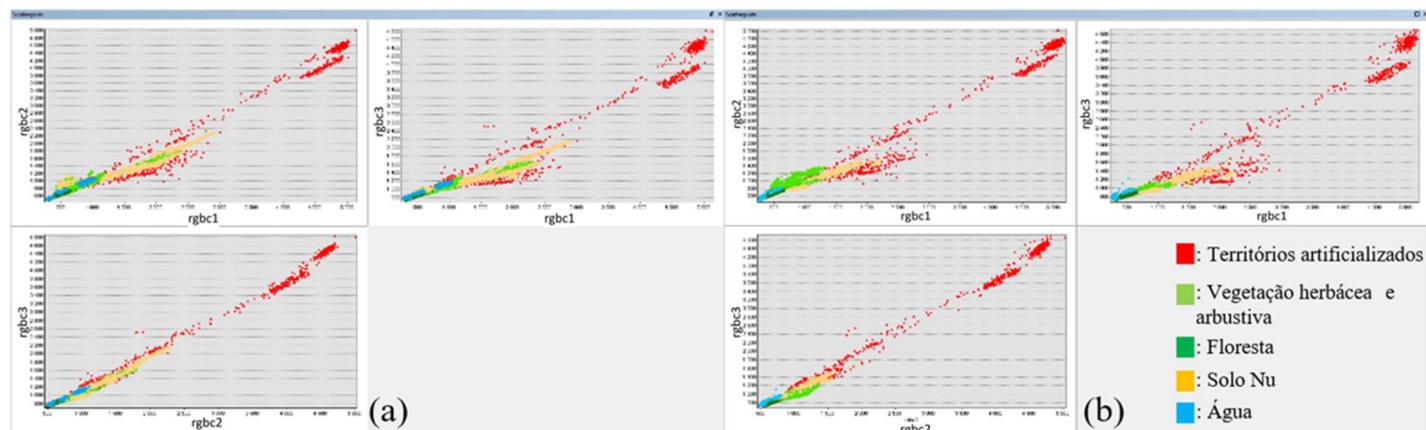


Figura 28 - Separabilidade das classes para a imagem de outono. (a) Antes da alteração dos polígonos (b) Depois da alteração dos polígonos

A **Figura 29** (a) apresenta uma visão geral da separabilidade das classes para a imagem de verão. Para a classificação feita com uma imagem recolhida neste período, as classes que se mostraram com menor separabilidade, foram as classes 1 e 2, as classes 1 e 4, as classes 2 e 4, as classes 2 e 5 e as classes 3 e 5. A **Figura 29** (b) mostra a separabilidade das classes depois da alteração dos polígonos.

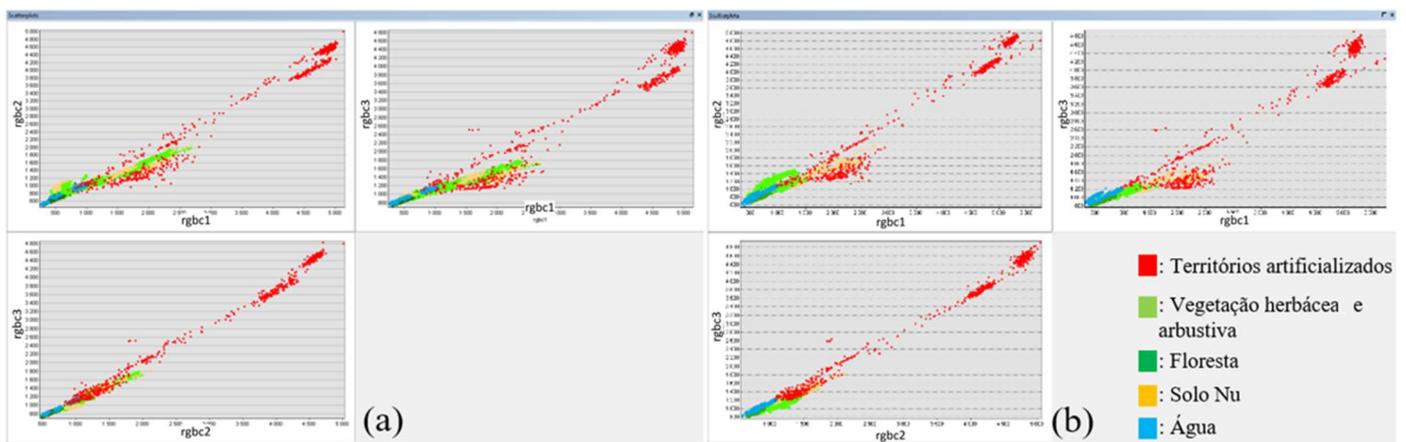


Figura 29 - Separabilidade das classes para a imagem de verão. (a) Antes da alteração dos polígonos (b) Depois da alteração dos polígonos

São notadas diferenças na separabilidade das classes, antes e depois da alteração dos dados de treino. Embora seja notado apenas visualmente, a sobreposição entre classes reduziu.

5.4. Mapas de ocupação do solo

Depois de obtidos os dados de treino considerados adequados para a classificação, foram criados os mapas de ocupação do solo com as classes anteriormente referidas. Após a criação dos mapas notou-se que uma zona do caso de estudo (mostrada na **Figura 30**) foi erroneamente classificada como água.

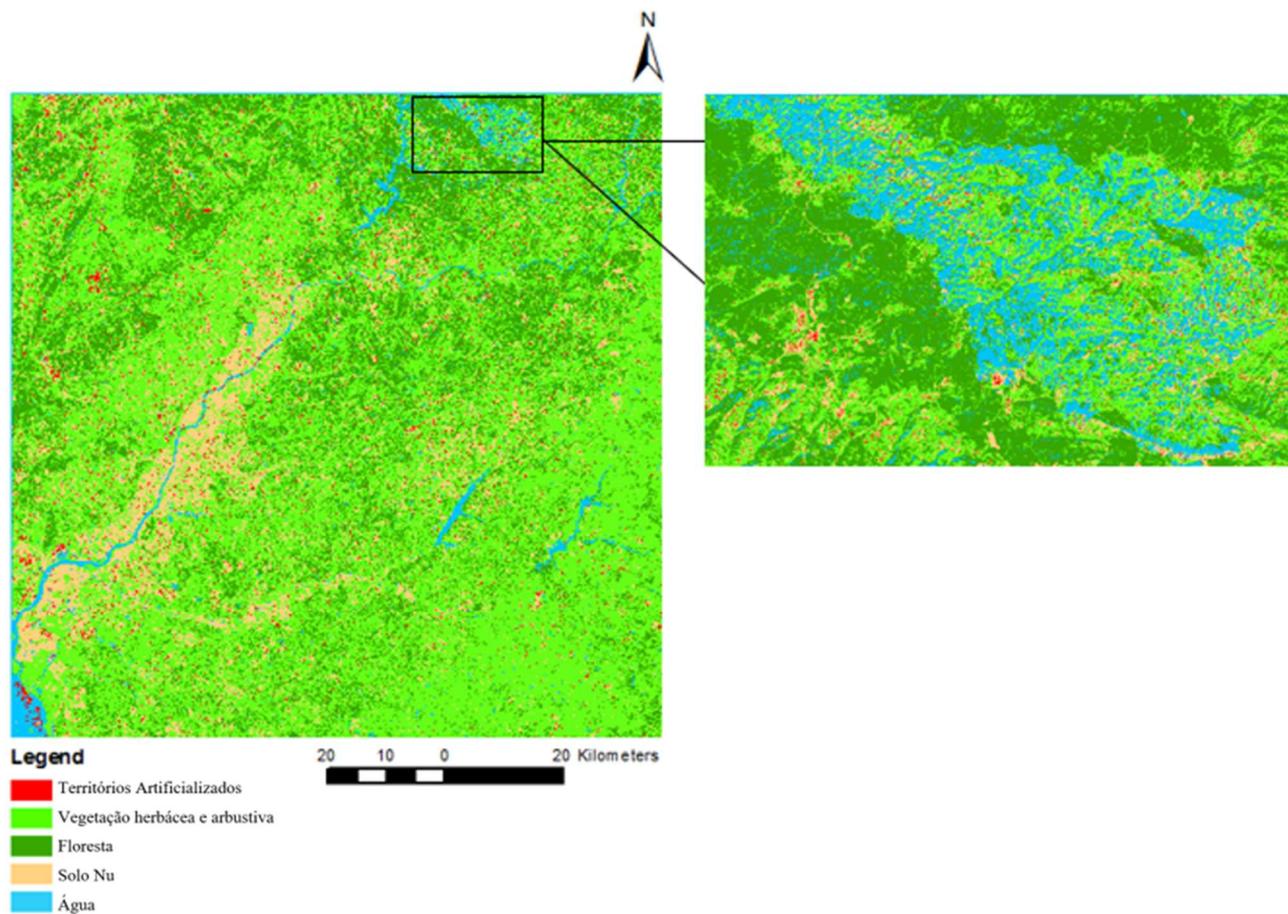


Figura 30 – Zona problemática, erroneamente classificada como água.

Inicialmente, pensou-se na hipótese do problema ser derivado a essa zona ser montanhosa e poder gerar sombras que seriam confundidas como água. Como a primeira imagem foi recolhida durante o Inverno (10/03/2020) o problema das sombras é ainda mais acentuado, uma vez que o sol está mais baixo. Foram classificadas as restantes imagens para o período de Outono (27/09/2019) e para o período de Verão (03/08/2019), visto que seria a altura do ano em que o sol estaria mais alto e por consequência haveria menos sombras.

No entanto, os resultados obtidos com estas imagens continuaram a incluir grandes zonas incorretamente classificadas como água. Foi feita uma análise da zona e concluiu-se que essa zona sofreu um incêndio, justificando assim a sua resposta espectral.

Na **Figura 31** são apresentados os mapas de ocupação do solo para o outono e para o verão.

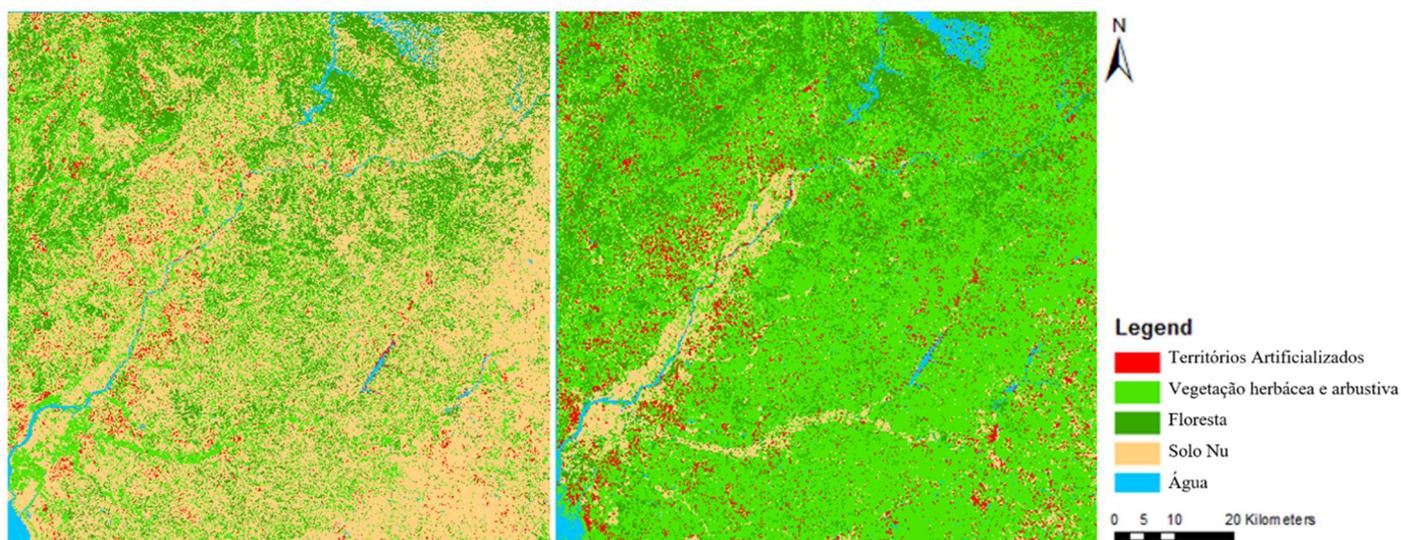


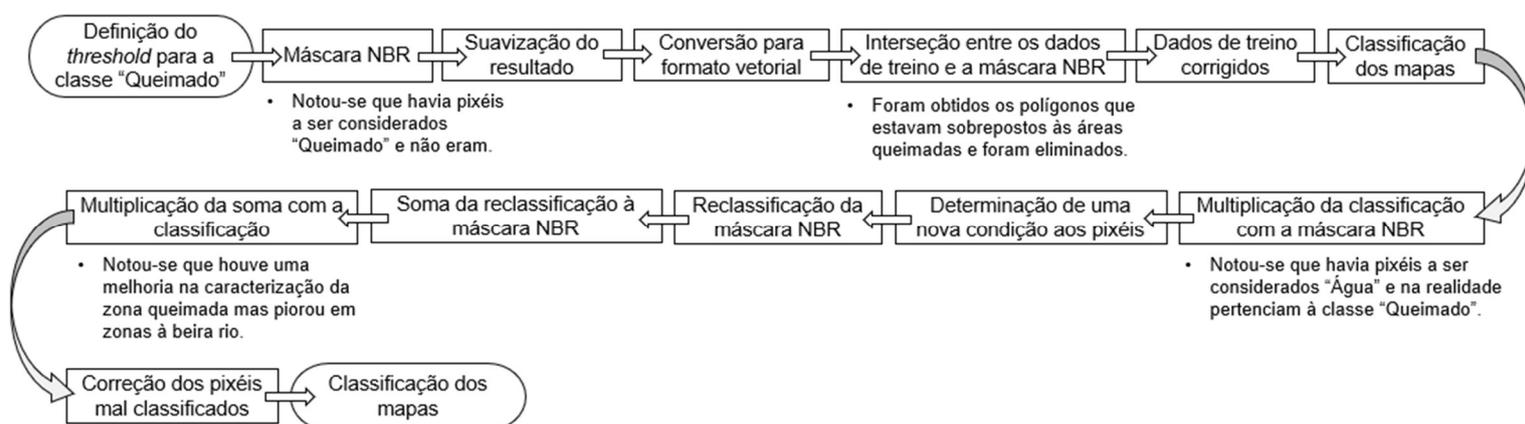
Figura 31 - Mapas de ocupação do solo para as imagens do período de outono (à esquerda) e de verão (à direita).

O principal enfoque deste trabalho é a floresta, e devido a restrições de tempo, estes mapas com 5 classes não foram validados. Estes foram utilizados para obter uma máscara de floresta (secção 5.6), que posteriormente foi validada.

5.5. Máscara NBR

Pensou-se numa metodologia que permitisse identificar zonas queimadas sem usar na classificação a classe “Queimado”, uma vez que o utilizador não é obrigado a saber *à priori* que a imagem que quer classificar tem áreas queimadas. Assim, a metodologia utilizada neste projeto passou por utilizar o NBR para a criação de uma máscara que permitisse a identificação de zonas queimadas, a utilização dessa máscara para a remoção de polígonos sobrepostos às áreas queimadas nos dados de treino e a reclassificação nos resultados da classificação dos pixéis como zonas queimadas. Todo este processo deu-se no ambiente ArcGIS. O processo é mostrado no

Fluxograma 2.



Fluxograma 2 – Processo para a classificação de mapas com zonas queimadas.

Para a criação de uma máscara com o NBR é importante saber entre que valores a classe “Água” e a classe “Queimado” variam no NBR (isto porque, são as classes que estão a ser confundidas), para perceber se os pixéis que estão a ser considerados como água e que na realidade são queimados podem ser separados. Notou-se que os valores do NBR variam muito entre as classes, sendo possível a separação destas. Os valores para a:

- Imagem de outono:
 - Valores da classe “Água”: $0,44 \leq NBR(x) \leq 0,75$
 - Valores da classe “Queimado”: $-0,3 \leq NBR(x) \leq 0$
- Imagem de verão:
 - Valores da classe “Água”: $0,30 \leq NBR(x) \leq 0,70$
 - Valores da classe “Queimado”: $-0,22 \leq NBR(x) \leq 0,22$

Sendo assim, para a imagem de verão os *thresholds* utilizados para que o pixel fosse considerado zona queimada foram: $-0,22 \leq NBR(x) \leq 0,22$ e para a imagem de outono foi: $-0,3 \leq NBR(x) \leq 0$.

Depois de criada a máscara NBR foram notados pixéis isolados que estavam a ser considerados queimados e na realidade não eram. Para suavizar o resultado foi utilizada a ferramenta “*Majority Filter*”, esta ferramenta caracteriza o pixel selecionado conforme a classe maioritária dos pixéis que o rodeiam, o número de pixéis vizinhos considerado foi 8. A ferramenta usada permite escolher o valor a atribuir ao pixel central da janela a considerar de duas formas:

- *Half* – Metade das células devem ter o mesmo valor e ser contíguas – por exemplo: em 8 vizinhos 4 têm de ter o mesmo valor,
- *Majority* – A maioria das células têm de ter o mesmo valor e ser contíguas – por exemplo: em 8 vizinhos 5 têm de ter o mesmo valor.

Ambos os métodos foram testados e avaliados visualmente, e os resultados mostraram que método *Half* foi o que permitiu uma maior suavização do resultado. O *output* obtido foi convertido para vetorial.

De forma a melhorar os dados de treino e para que outras classes não sejam confundidas com o queimado, foi feita a interseção entre os dados de treino e as zonas queimadas. Com esta operação foram obtidos os polígonos sobrepostos às áreas queimadas, estes foram selecionados e eliminados dos dados de treino.

Com os dados de treino corrigidos foram feitas novas classificações dos mapas de ocupação do solo na plataforma F-TEP. A classificação obtida é multiplicada pela máscara NBR em formato *raster*. O resultado é mostrado na **Figura 32** e na **Figura 33**, para a imagem de verão e de outono, respectivamente.

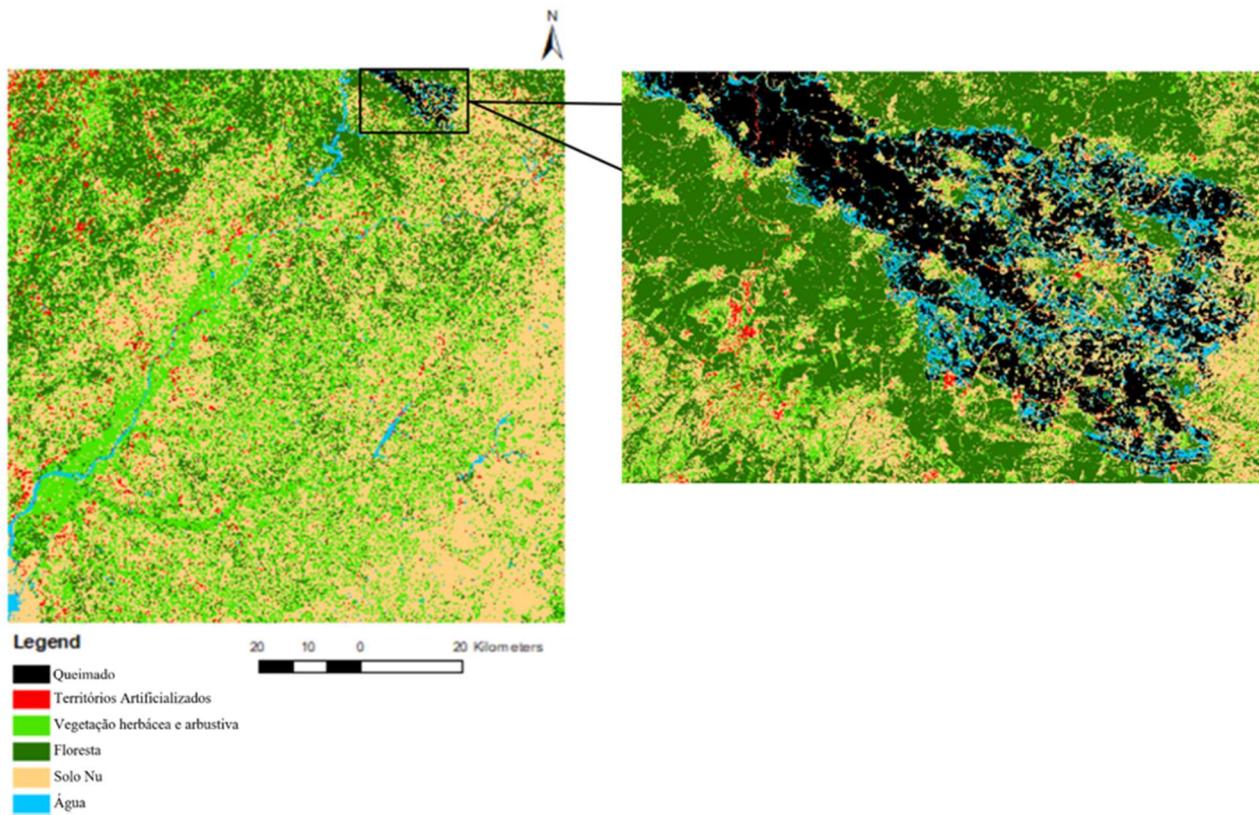


Figura 32 - Classificação da imagem de verão já com a classe "Queimado".

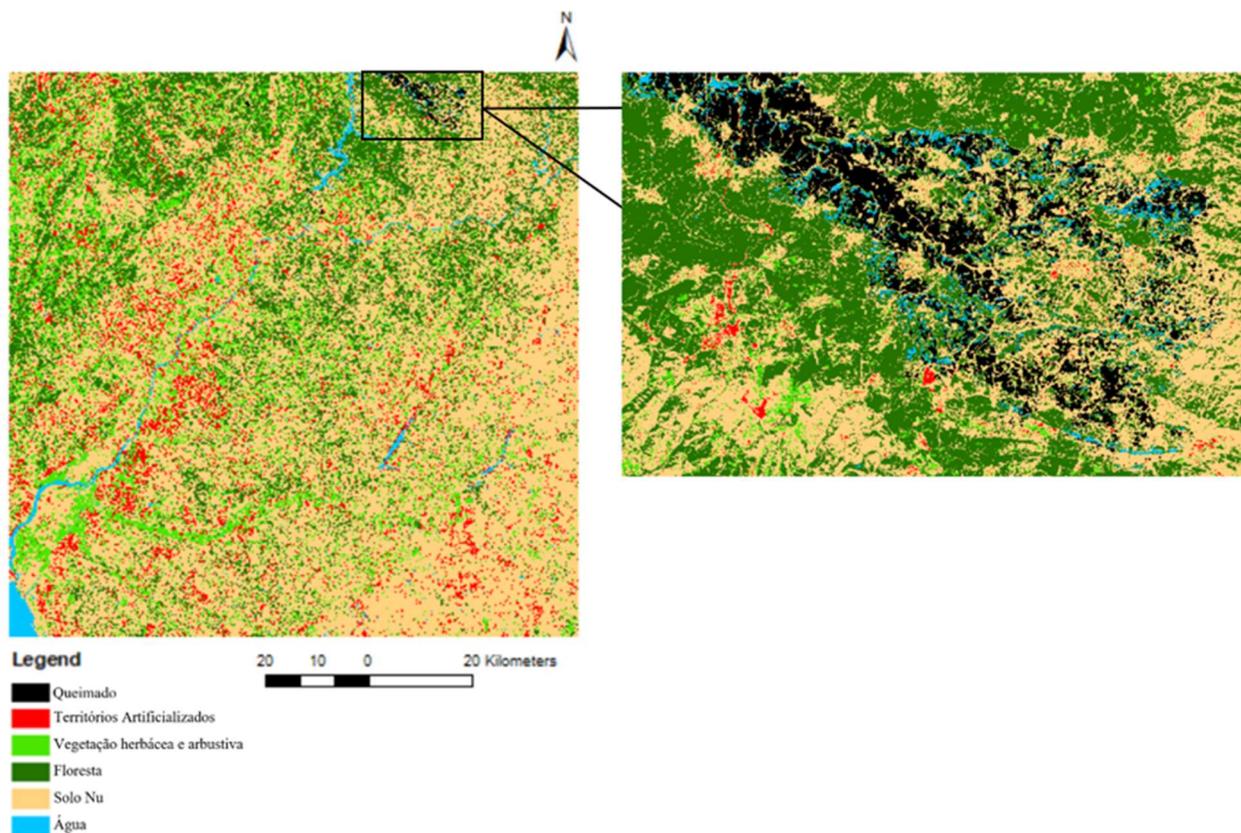


Figura 33 - Classificação da imagem de outono já com a classe "Queimado".

Mesmo com a máscara NBR a zona problemática continuou a apresentar pixels classificados erroneamente. Assim sendo, foi imposta uma nova condição aos pixels que seriam considerados queimados. Para que isto fosse possível começou-se por fazer uma reclassificação da classificação obtida anteriormente, isto para que pudéssemos ter uma classe para a água e outra para o que não é água.

Para a classificação de outono, os pixels pertencentes à classe “Água” com valores entre: $-0,3 \leq NBR(x) \leq 0,3$ foram também considerados queimados. Para a classificação de verão o mesmo acontece com os pixels com valores entre: $-0,25 \leq NBR(x) \leq 0,25$.

Com a condição imposta nos pixels, foi feita uma reclassificação à máscara NBR, em que as classes seriam:

- Classe 0 = Água,
- Classe 1 = Queimado,
- Classe 2 = *NoData*.

Para ter uma máscara com todos os pixels que pertencem à classe “Queimado”, a reclassificação foi somada à máscara NBR. Com a soma foram obtidos os valores: “0” para os pixels que são “Água” nas duas máscaras, “1” para os pixels que são “Água” numa máscara e “Queimado” noutra, “2” para os pixels que são “Queimado” nas duas máscaras e “3” para os pixels

que são “Queimado” numa máscara e “NoData” noutra. O *output* da soma entre as duas máscaras é reclassificado em:

- Classe 0 = Queimado,
- Classe 1 = Não Queimado.

Por fim, a soma entre a reclassificação e a máscara NBR é multiplicada com a classificação. Sendo que na classificação a classe 0 é a classe “Queimado” e as restantes classes permanecem iguais.

Com a nova condição imposta sobre os pixéis, houve uma melhoria na zona queimada, no entanto, piorou em zonas à beira rio. Os pixéis que se encontravam perto de água estavam a ser considerados zonas queimadas. Isto, foi mais saliente no período de Verão, uma vez que o nível da água é mais baixo e consequentemente há mais areias à superfície, o que fez com que a classe “Solo Nu” fosse confundida com a classe queimada. Como pode ser visto na **Figura 34** e na **Figura 35**.

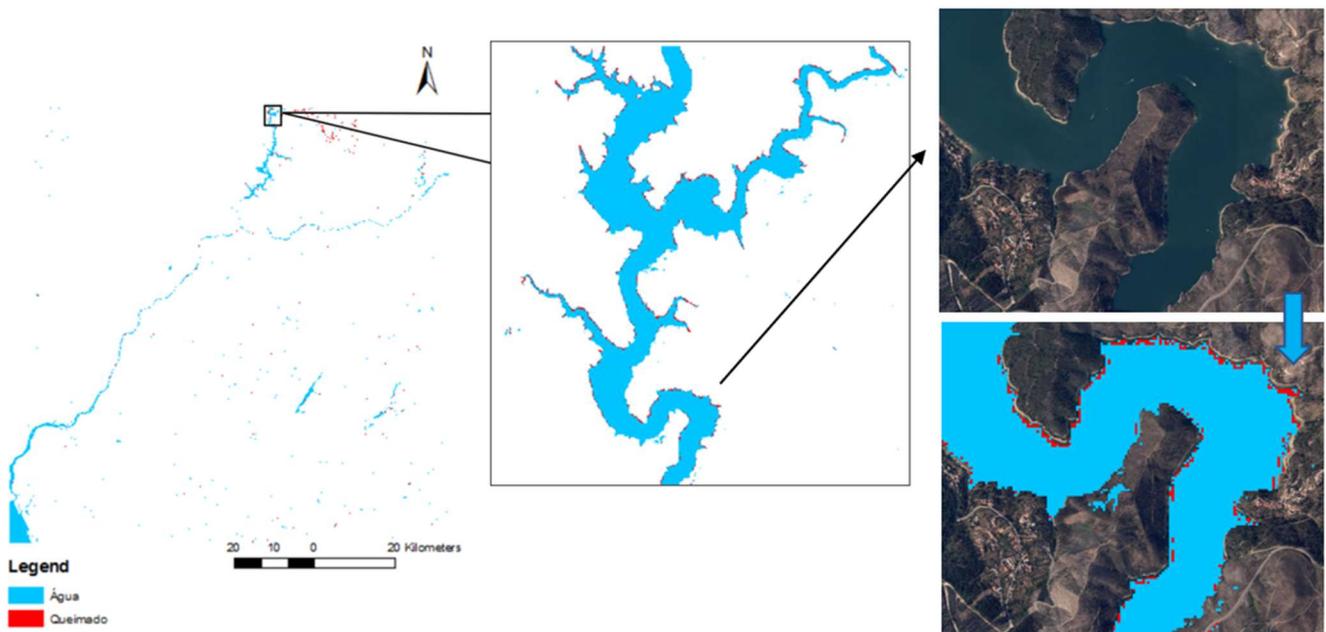


Figura 34 - Confusão entre as classes "Solo Nu" e "Queimado" na imagem de outono.

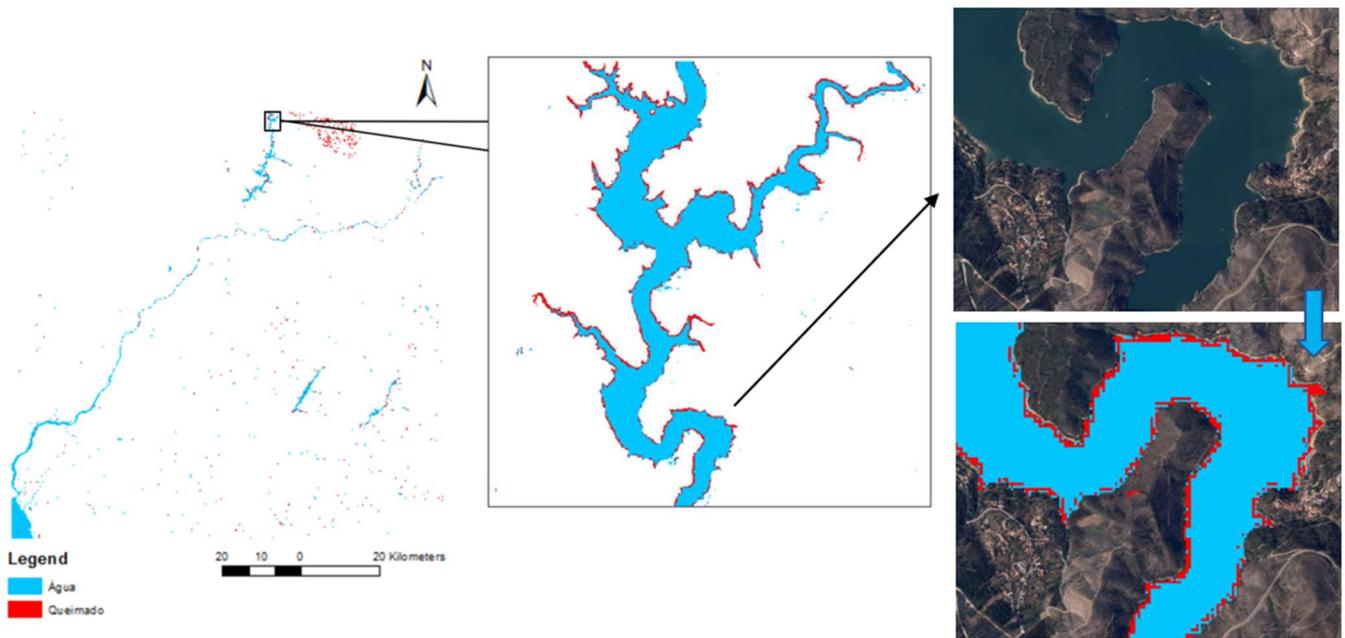


Figura 35 - Confusão entre as classes "Solo Nu" e "Queimado" na imagem de verão.

Para corrigir os pixels que estão a ser mal classificados foi utilizada a ferramenta “*Block Statistics*”. Esta ferramenta permite utilizar uma janela maior que na ferramenta “*Majority Filter*”, ou seja, permite considerar um maior número de vizinhos. Aquando o teste da melhor janela a ser utilizada, a janela 7 por 7 pixels foi a que visualmente mostrou ter melhor desempenho.

Posteriormente, multiplicou-se o *output* obtido pela ferramenta “*Block Statistics*” com a classificação. As zonas perto do rio deixaram de ser um problema, a qualidade destes resultados foi apenas validada visualmente. Os resultados obtidos são mostrados na **Figura 36** e na **Figura 37**.

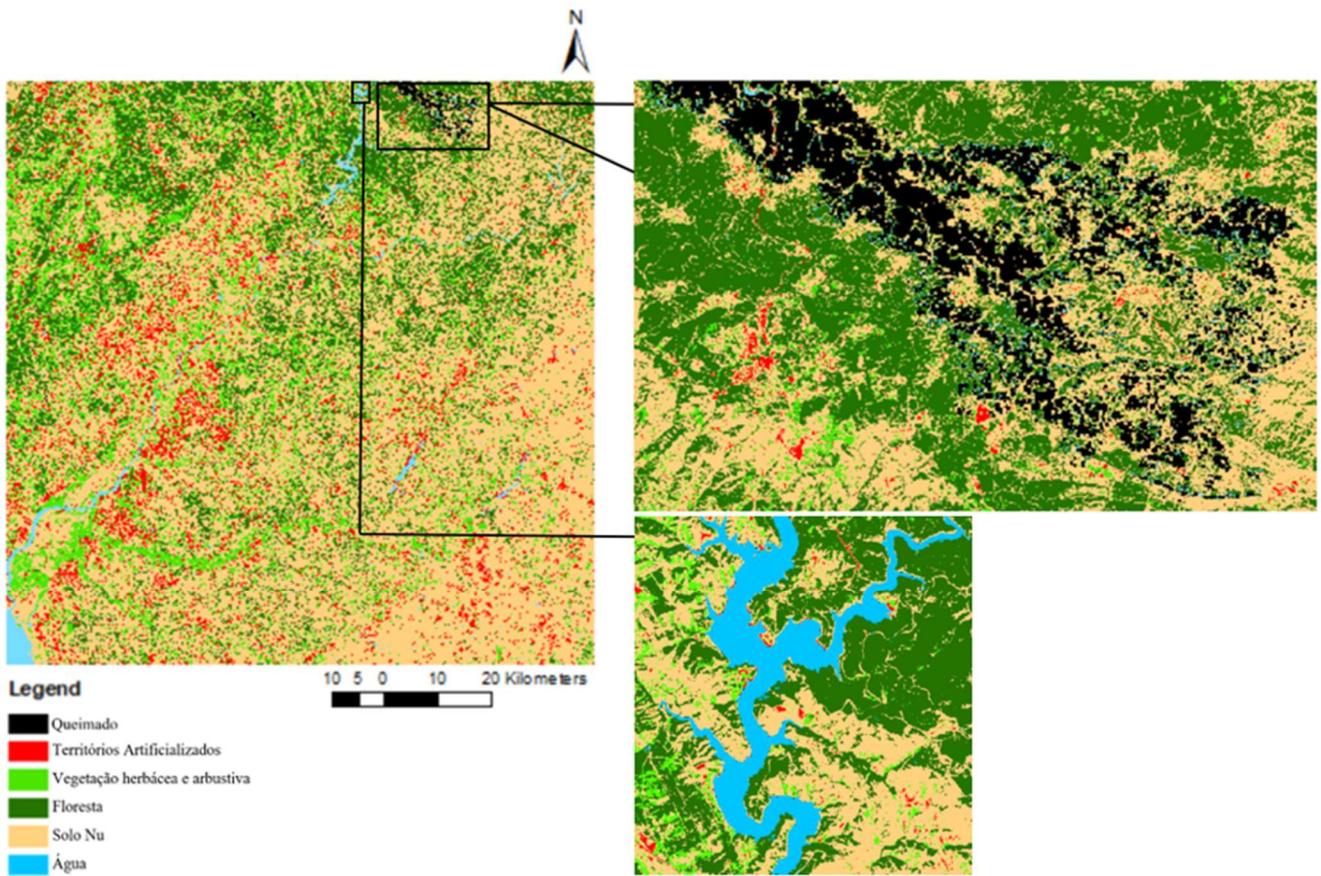


Figura 36 - Classificação final da imagem de outono.

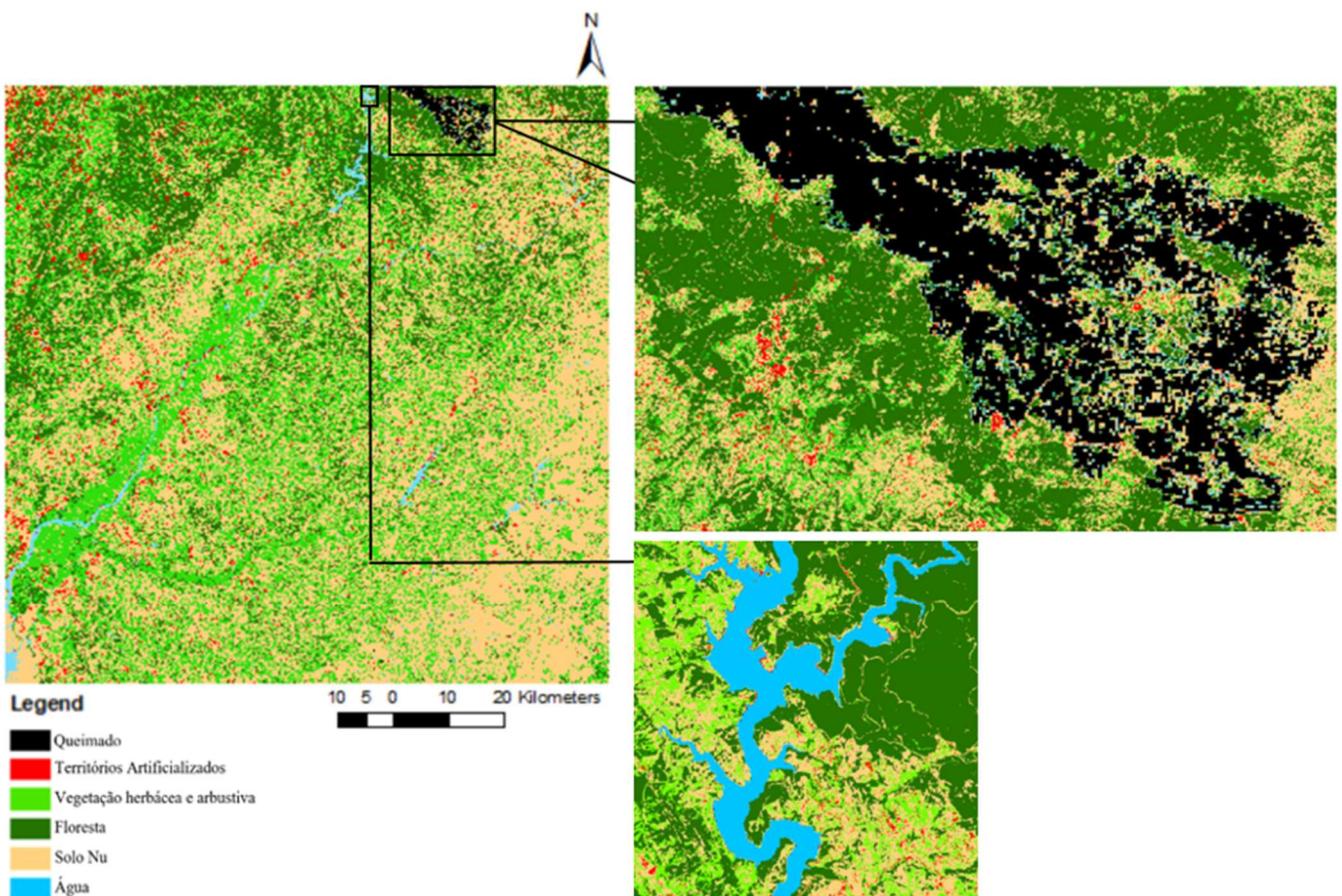


Figura 37 - Classificação final da imagem de verão.

5.6. Máscara da floresta

Tendo a classificação final dos mapas de ocupação do solo para cada uma das imagens, a máscara da floresta é obtida pela simplificação dos mapas em duas classes: a classe “Floresta” e a classe “Não Floresta”, onde:

- Classe 0 = Não Floresta,
- Classe 1 = Floresta.

Na **Figura 38** é mostrada a máscara de floresta para o período de inverno, na **Figura 39** para o período de outono e na **Figura 40** para o período de verão.

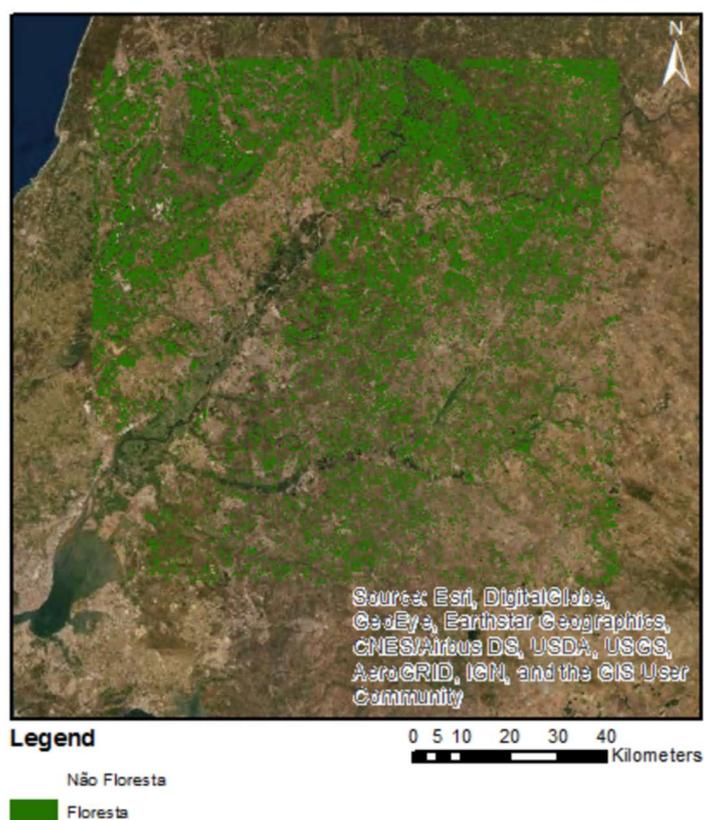


Figura 38 – Máscara da floresta para o período de inverno.

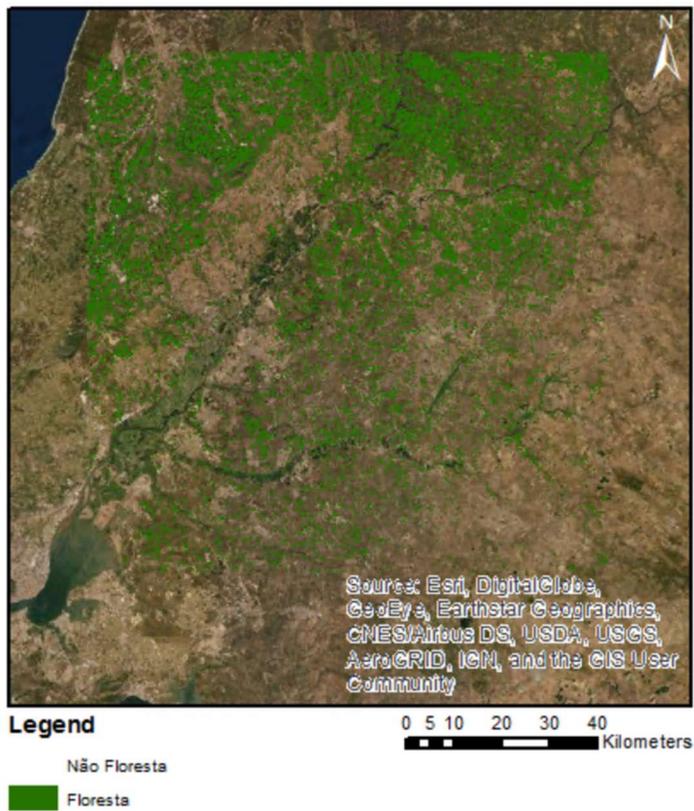


Figura 39 - Máscara da floresta para o período de outono.



Figura 40 - Máscara da floresta para o período de verão.

5.6.1. Validação da máscara da floresta

Depois de obtidas as máscaras de floresta, foi feita a validação da máscara do período de inverno e da máscara do período de verão. O RAIZ não dispunha de dados de referência para a validação da máscara de outono e devido às limitações de tempo não foram criados dados de referência. Assim sendo, a máscara de outono não foi validada.

Para a validação da máscara de floresta do período de inverno foram utilizados os dados de referência do inventário florestal na zona RAIZ3 (recolhidos a 12/02/2020). Para a validação da máscara de floresta do período de verão foram utilizados os dados de referência da validação da máscara de floresta (recolhidos a 03/2019) na zona RAIZ3.

Para ser feita a validação precisamos da “verdade do terreno” e o que foi obtido na classificação, foram por isso, adicionadas duas colunas com essa informação às tabelas dos dados de referência. A coluna relativa à “verdade do terreno” foi preenchida com base nos dados fornecidos pela RAIZ: o campo é todo preenchido com “1”, são selecionados os valores “*Non Forest*” e preenchidos com “0”. A coluna relativa aos dados da classificação foi preenchida com apoio da ferramenta “*Extract Values to Points*”, que permite obter o valor do *raster* a que o ponto, que está a ser validado, está sobreposto fornecendo-lhe esse valor.

Depois de preenchidas as duas novas colunas foi calculada a matriz de confusão pela ferramenta “*Compute Confusion Matrix*” disponível no *software* MatLab, em que são obtidos os valores previstos nas colunas e os valores reais nas linhas. As matrizes de confusão obtidas são mostradas na **Tabela 3** e na **Tabela 4**.

		Valores Previstos		Total	Exatidão do produtor (%)
		0	1		
Valores Reais	0	34	1	35	97
	1	44	43	87	51
Total		78	44	122	
Exatidão do utilizador (%)		44	98		

Tabela 3 - Matriz de confusão para a classificação de inverno.

		Valores Previstos		Total	Exatidão do produtor (%)
		0	1		
Valores Reais	0	95	5	100	0,950
	1	218	220	438	0,502
Total		313	225	538	
Exatidão do utilizador (%)		30	98		

Tabela 4 - Matriz de confusão para a classificação de verão.

Para a classificação de inverno, a exatidão do utilizador na classe 0 foi de 44% e na classe 1 foi de 98%, a exatidão do produtor para a classe 0 foi de 97% e para a classe 1 foi de 51%. Para a classificação de verão, a exatidão do utilizador na classe 0 foi de 30% e na classe 1 foi de 98%, a exatidão do produtor para a classe 0 foi de 95% e para a classe 1 foi de 50%.

Para o verão a exatidão global foi de 58% e para o inverno de 63%. Tanto para a máscara de floresta do verão como para a máscara de floresta do inverno o problema deu-se nos FN, ou seja, os pontos que estão classificados como classe “Não Floresta” e na realidade pertencem à classe “Floresta”. Isto deve-se essencialmente a três razões, aos dados estarem em zonas:

- Perto de caminhos rurais, sendo por isso que maior parte dos pontos são confundidos ou pela classe “Vegetação herbácea e arbustiva”, quando estão perto de caminhos rurais estreitos, ou pela classe de “Solo Nu”, quando estão perto de caminhos rurais largos, como pode ser visto na **Figura 41**. São erros provocados pela exatidão posicional e pela resolução limitada das imagens.



Figura 41 - Pontos perto de caminhos rurais. Na primeira linha de imagens temos como *base layer* o *basemap* do ArcMap, na segunda linha temos como *base layer* a classificação de verão, sendo que a verde claro temos a classe “Vegetação herbácea e arbustiva”, a verde escuro a classe “Floresta” e a castanho claro a classe “Solo Nu”.

- De plantação jovem, sendo confundidos pelo classificador como “Vegetação herbácea e arbustiva” e “Solo Nu”, como pode ser visto na **Figura 42**.

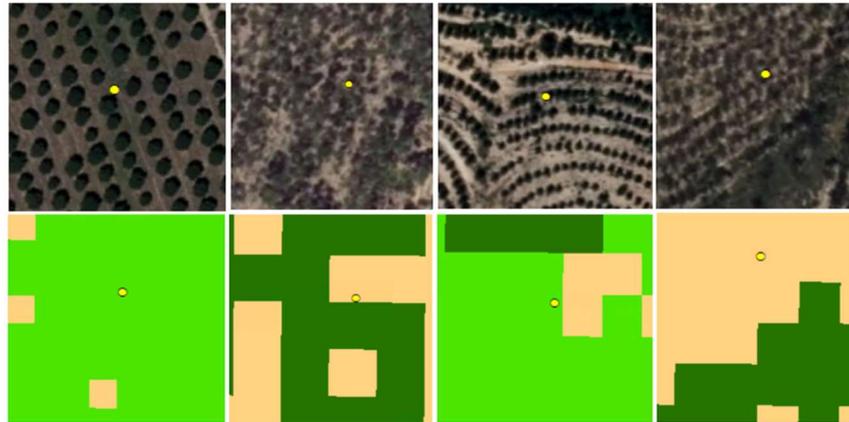


Figura 42 - Pontos em plantações jovens. Na primeira linha de imagens temos como *base layer* o *basemap* do ArcMap, na segunda linha temos como *base layer* a classificação de verão, sendo que a verde claro temos a classe “Vegetação herbácea e arbustiva”, a verde escuro a classe “Floresta” e a castanho claro a classe “Solo Nu”.

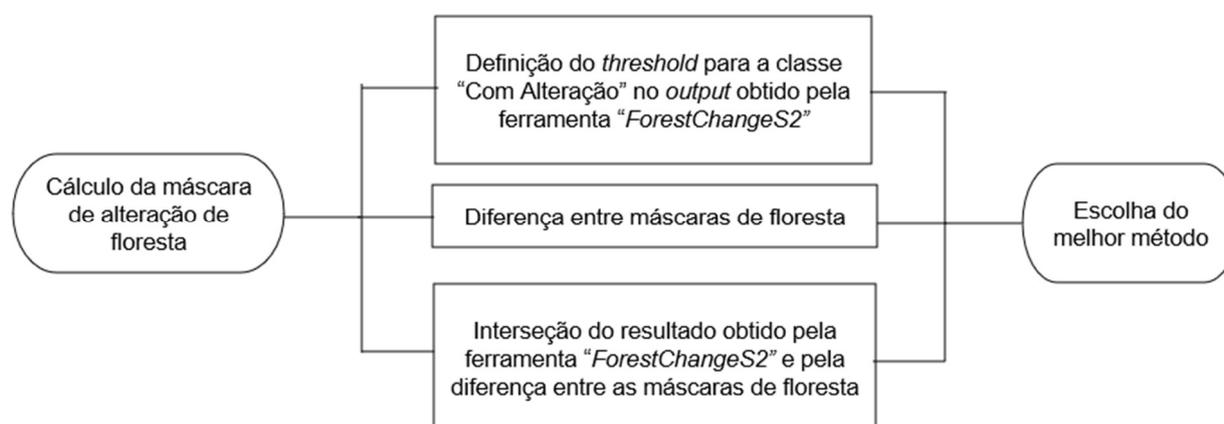
- Em que a floresta é dispersa e apresenta vegetação no solo, sendo confundido com a classe “Vegetação herbácea e arbustiva” ou no caso da floresta ser muito dispersa e ser confundida com a classe do “Solo Nu”, como pode ser visto na **Figura 43**.



Figura 43 - Pontos em floresta dispersa e com vegetação e solo nu. Na primeira linha de imagens temos como *base layer* o *basemap* do ArcMap, na segunda linha temos como *base layer* a classificação de verão, sendo que a verde claro temos a classe “Vegetação herbácea e arbustiva”, a verde escuro a classe “Floresta” e a castanho claro a classe “Solo Nu”.

5.7. Alterações na classe floresta

O processo recorrido para a escolha do melhor método para a identificação de zonas de alteração na floresta é mostrado no **Fluxograma 3**.



Fluxograma 3 – Processo para a escolha do melhor método para a identificação de zonas de alteração na floresta

As alterações ocorridas na floresta entre o outono e inverno, entre o verão e o outono e entre o verão e o inverno foram calculadas com base em três métodos:

- Pela ferramenta “*ForestChangeS2*” existente na Plataforma F-Tep, sendo esta calculada com o NDVI;
- Pela diferença entre máscaras de floresta, subtraindo à máscara de floresta do tempo final a máscara de floresta do tempo inicial;
- Pela interseção dos dois métodos referidos anteriormente.

Excluindo o primeiro método referido anteriormente, todo o restante processo deu-se no ambiente ArcGIS.

Foram então definidos os *thresholds* no *output* obtido pela ferramenta “*ForestChangeS2*”. Foi feita uma reclassificação à máscara obtida e foi definido entre que valores seria considerada a classe “Sem Alteração” e a classe “Com Alteração”, entre os períodos:

- Outono – inverno:
 - 8-17000: “Sem Alteração”
 - 17000-32767: “Com Alteração”
- Verão – outono:
 - 18-16000: “Sem Alteração”
 - 16000-32767: “Com Alteração”
- Verão – inverno:
 - 7-12700: “Sem Alteração”

- 12700-32767: “Com Alteração”

Foi calculada a diferença de máscaras de floresta (sendo que os pixéis com floresta tinham o valor “1” e os sem floresta o valor “0”) entre os períodos, como pode ser visto pela **Equação 6**, **Equação 7** e **Equação 8**:

- Outono – inverno:

$$diferença = máscara\ de\ floresta_{inverno} - máscara\ de\ floresta_{outono} \quad (6)$$

- Verão – outono:

$$diferença = máscara\ de\ floresta_{outono} - máscara\ de\ floresta_{verão} \quad (7)$$

- Verão – inverno:

$$diferença = máscara\ de\ floresta_{inverno} - máscara\ de\ floresta_{verão} \quad (8)$$

Os pixéis que na diferença obtêm o valor “0” significa que não apresentam alteração, uma vez que o pixel é “Floresta” em ambos os períodos ou “Não Floresta”. Os pixéis que obtêm o valor “1” significa que apresentam alteração, sendo que no período inicial esse pixel é “Floresta” e no final é “Não Floresta”. Os pixéis que obtêm o valor “-1” significa que apresentam alteração, sendo que no período inicial é “Não Floresta” e no final é “Floresta”.

Para ter apenas duas classes, uma classe com a alteração e outra sem alteração foi feita uma reclassificação da diferença entre máscaras, obtendo para todos os períodos as classes:

- Classe 0 – Sem Alteração
- Classe 1 – Com Alteração

Por fim, foram intersetados os dois métodos anteriormente referidos nos períodos, como pode ser visto pela **Equação 9**, **Equação 10** e **Equação 11**:

- Outono – inverno:

$$interseção_{outono - inverno} = NDVI_{outono - inverno} \cap diferença_{outono - inverno} \quad (9)$$

- Verão – outono:

$$interseção_{verão - outono} = NDVI_{verão - outono} \cap diferença_{verão - outono} \quad (10)$$

- Verão – inverno:

$$interseção_{\text{verão} - \text{inverno}} = NDVI_{\text{verão} - \text{inverno}} \cap diferença_{\text{verão} - \text{inverno}} \quad (11)$$

Com a interseção são obtidos os pixels com alteração, tanto no método utilizando o NDVI como pela diferença entre máscaras, com o valor “1”. Para o pixel ser considerado alteração tem de ser alteração nos dois métodos, caso contrário não é considerado alteração.

Foi analisado visualmente o melhor método e o que se mostrou mais competente foi a interseção entre ambos. A validação é apenas feita para o método mais competente, como pode ser visto na secção 5.6.1.

5.7.1. Validação da máscara de alterações da floresta

A validação foi feita ao nível do polígono, isto porque, uma alteração na floresta não ocorre ao nível do pixel, o mais comum é essas alterações serem zonas de corte e caso fosse considerado pixel a pixel, uma vez que estes são seleccionados aleatoriamente, dois ou mais pontos poderiam ocorrer na mesma zona, e nesse caso, os resultados não seriam fidedignos.

Utilizando a ferramenta “*Region Group*” todas as regiões existentes na interseção são identificadas unicamente, sendo assim possível validá-las uma a uma. Seleccionaram-se todas as regiões que têm até só 4 pixels porque muito dificilmente 4 pixels formam uma floresta, então não faz sentido que esses sejam validados. Foi feita uma reclassificação dos dados, onde:

- *NoData* = Sem alteração,
- Classe 1 = Com alteração.

O ficheiro *raster* obtido foi convertido para vetorial. Com a ferramenta “*Feature to Point*”, foi criado um centroide em cada um dos polígonos.

Alguns exemplos de mudanças na classe floresta podem ser vistos na **Figura 44**, na **Figura 45** e na **Figura 46**. A **Figura 44** diz respeito à alteração da floresta entre os períodos verão e inverno, **Figura 45** diz respeito à alteração da floresta entre os períodos outono e inverno e a **Figura 46** diz respeito à alteração da floresta entre os períodos verão e outono, na área que é representada.

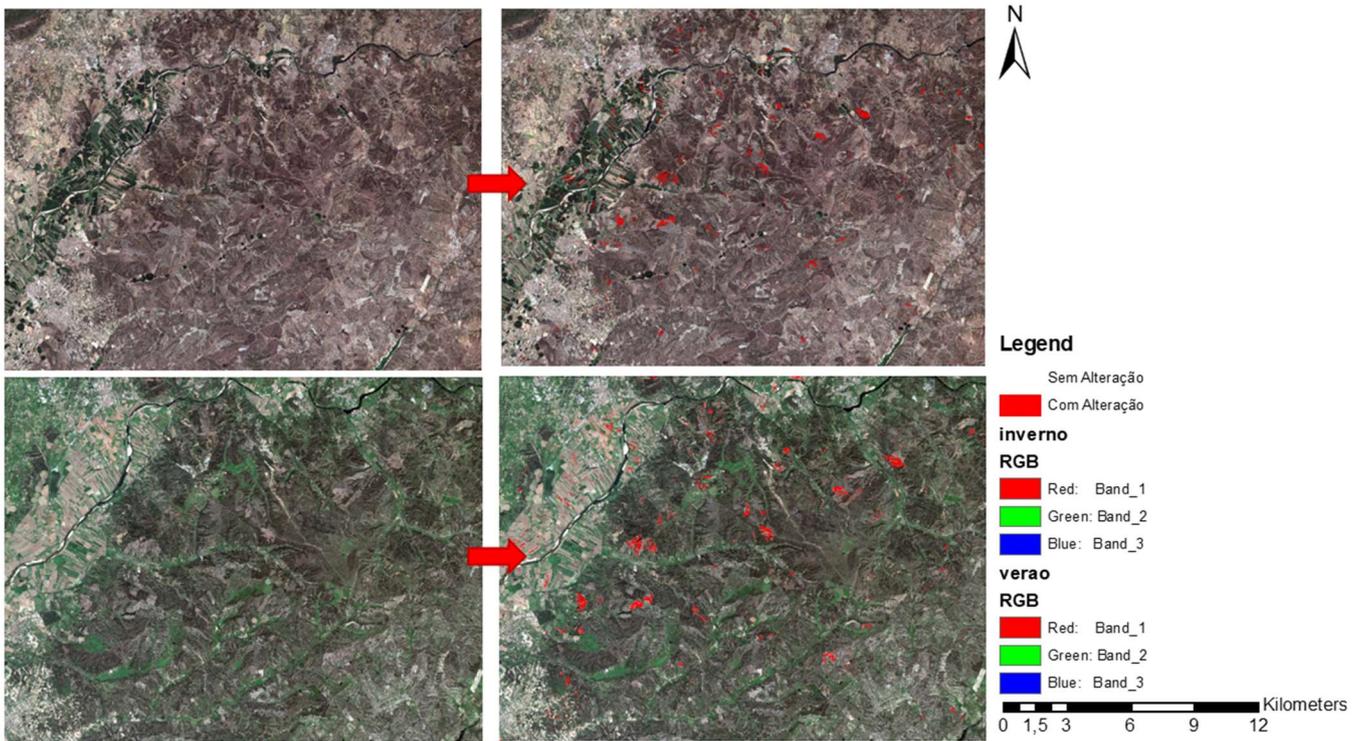


Figura 44 - Zonas de alteração entre o verão e o inverno. Em cima a imagem de verão seguida das alterações e por baixo a imagem de inverno seguida das alterações.

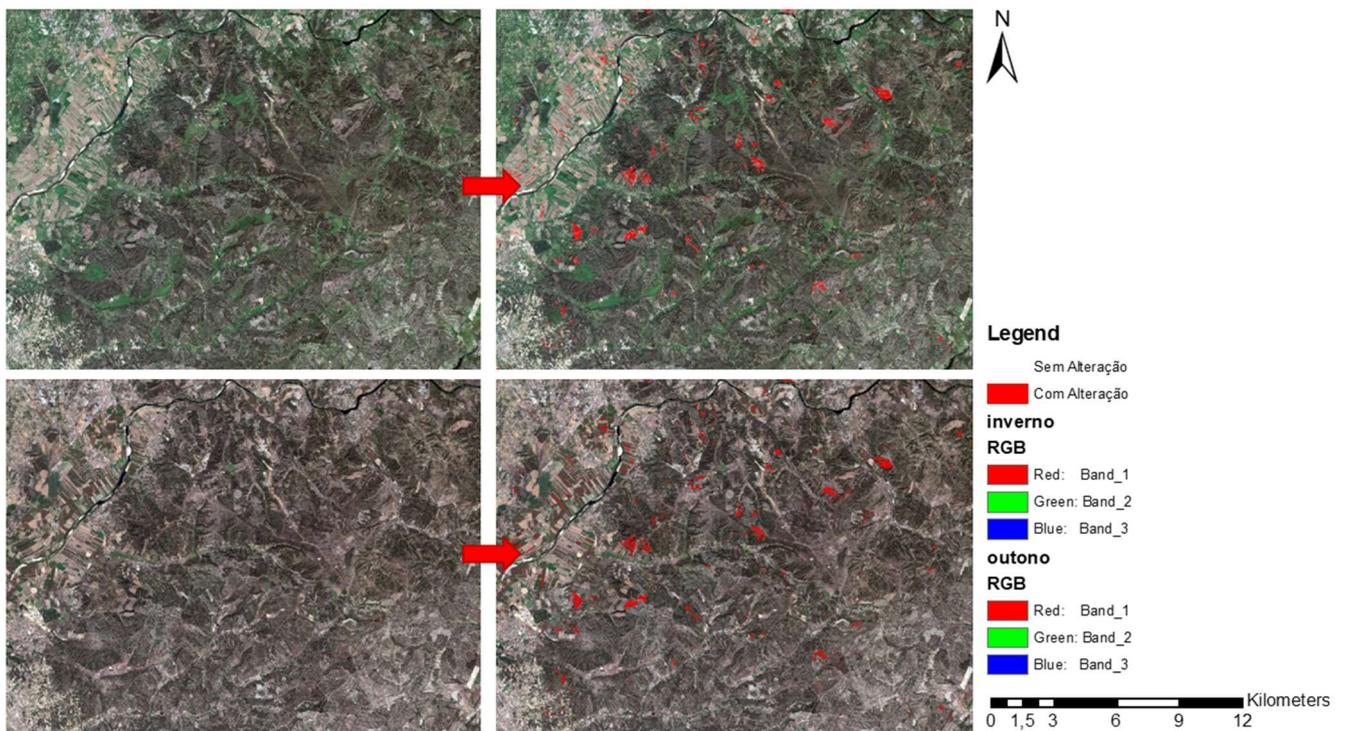


Figura 45 - Zonas de alteração entre o outono e o inverno. Em cima a imagem de inverno seguida das alterações e por baixo a imagem de outono seguida das alterações.

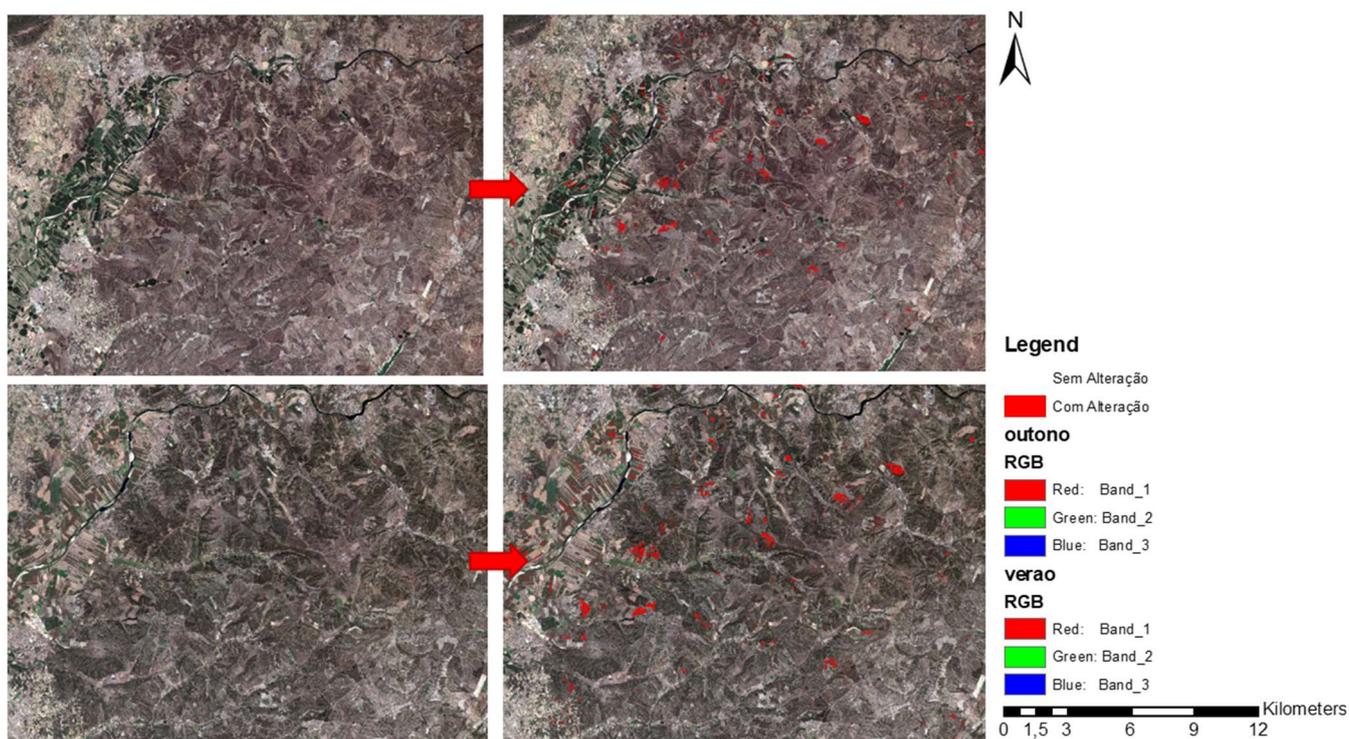


Figura 46 – Zonas de alteração entre o verão e o outono. Em cima a imagem de verão seguida das alterações e por baixo a imagem de outono seguida das alterações.

Para se fazer a validação das regiões com alteração de floresta foram selecionadas, aleatoriamente, 150 regiões. Essas regiões foram exportadas para uma nova tabela, onde foram adicionadas duas colunas, uma com os dados de referência (o campo foi todo preenchido por “1”) e outra com a classificação (preenchida com base na análise de imagens de satélite). O tipo de amostragem utilizada para a validação não permite identificar omissões, isto é, como só é validado o que é alteração de floresta não podemos saber o desempenho da máscara para zonas que não foram identificadas como alteradas.

Para a alteração de floresta no período outono-inverno, a exatidão do utilizador na classe 1 foi de 63%. Para a alteração de floresta no período verão-inverno, a exatidão do utilizador na classe 1 foi de 57%. Para a alteração de floresta no período verão-outono, a exatidão do utilizador na classe 1 foi de 41%.

As alterações de floresta ocorridas entre o outono e o inverno tiveram uma exatidão global de 63%, entre o verão e o inverno foi de 57% e entre o verão e o outono foi de 41%. Os erros obtidos deveram-se essencialmente à falta de dados de treino na classe que estava a ser erroneamente classificada como “Floresta”.

6. Discussão & Conclusões

Com este projeto, pretendeu-se testar uma metodologia que o RAIZ pudesse utilizar para a criação de uma máscara de floresta e de uma máscara de alterações da floresta. Para isso, foram identificadas as plataformas que disponibilizam informação sobre a cobertura do solo, potencialmente úteis para a monitorização da floresta. Foi analisada a informação disponível nas várias plataformas e as suas características. As plataformas que se mostraram mais promissoras foram:

- *Sentinel Hub*,
- *Terramonitor*,
- *Giovanni*,
- *Land Cover*,
- *Global Forest*,
- F-Tep.

Foi feita uma análise sobre os produtos existentes nestas plataformas e notou-se que os seguintes portais não se mostraram capazes de concretizar os objetivos, isto porque, para o portal:

- *Sentinel Hub*: Os produtos mapas de agricultura e mapas com o índice de vegetação apresentam uma resolução máxima de 100m;
- *Terramonitor*: Os produtos mapas com o índice VESA, mapas de vegetação e mapas do serviço *ForestMonitor* são pagos e tomou-se como prioridade o teste de produtos gratuitos;
- *Giovanni*: Os produtos mapas com o LAI e mapas com a fração verde apresentam um período de atualização mensal e uma resolução máxima de 75km;
- *Land Cover*: Os produtos mapas de floresta aberta e fechada e mapas com os tipos de floresta apresentam apenas dados até 2015 e uma resolução máxima de 10km;
- *Global Forest*: Os produtos mapas de cobertura arbórea e mapas de cobertura do solo apresentam uma resolução espacial de 300m e 30, respetivamente, e só existem dados até ao ano 2015;

No portal F-Tep, os produtos disponíveis de interesse para o RAIZ são mapas da alteração de floresta, mapas de cobertura do solo e mapas com índices de vegetação. Estes produtos, apresentam dados com resolução espacial máxima de 10m e existem dados até ao presente, sendo

atualizados de 5 em 5 dias ou de 10 em 10 dias. Com estas características e as limitações referidas anteriormente, conclui-se que é o único portal, dos estudados, capaz de alcançar os objetivos propostos.

Relativamente às outras plataformas, a plataforma F-Tep apresenta uma interface bastante amigável e de fácil entendimento, mesmo para quem não seja da área em específico. Possui diversas ferramentas com resolução capaz de realizar um estudo sobre a floresta e das ferramentas testadas, estas apresentaram bons resultados. Em contrapartida, o acesso a estas ferramentas é limitado a 100 *coins*, cada processo custa 1 *coin*.

Como afirmado na introdução, esta tese tinha como objetivo principal a identificação automática do que é floresta e não é floresta e a identificação automática de zonas de alteração de floresta, estes objetivos foram concretizados pelas seguintes fases:

- 1) Criação do mapa de uso do solo,
- 2) Criação da máscara NBR,
- 3) Multiplicação da máscara NBR com o mapa de uso do solo criado anteriormente,
- 4) Simplificação do mapa final em 2 classes: a classe “Floresta” e a classe “Não Floresta”,
- 5) Com a simplificação do mapa foram calculadas as alterações na classe da floresta pela ferramenta “*ForestChangeS2*”, pela diferença entre os períodos e pela interseção de ambos os métodos.

Como objetivos secundários, foi dito que seriam criados mapas de ocupação do solo e seriam validadas as máscaras de floresta para o período de inverno e de verão e as máscaras de alterações da floresta.

Todos os objetivos foram cumpridos com sucesso. Os resultados obtidos são explicados e discutidos nas secções 6.1 e 6.2.

6.1. Identificação automática de florestas

A identificação automática de florestas passou pelos seguintes passos:

- Escolha do caso de estudo,
- Criação dos dados de treino,
- Mapas de ocupação do solo,
- Máscara da floresta.

Inicialmente foi escolhido o caso de estudo, de entre as quatro áreas de interesse do RAIZ foi escolhida a área RAIZ3. Foram criados dados de treino na área de interesse escolhida e foi analisada a separabilidade entre classes e posteriormente foram eliminados os polígonos (dos dados de treino) que estavam a gerar confusão entre classes, no ambiente ArcGIS. Ao fazer essa seleção e exclusão de polígonos, estão a ser escolhidas amostras de treino que podem não ser representativas pois apenas foram escolhidos os polígonos que contribuísssem para a separabilidade das classes. Nesta análise notou-se que as classes que geraram mais confusão entre si foram: as classes “Solo Nu” e “Territórios Artificializados” e entre as classes “Água” e “Floresta”, já era o esperado uma vez que estas classes apresentam respostas espectrais semelhantes.

Já com os dados de treino criados, foram classificadas as imagens com base nos mesmos, obtendo os mapas de ocupação do solo. Na criação destes mapas foram notadas zonas queimadas e, como o utilizador não tem de saber *à priori* que a imagem que quer classificar tem áreas queimadas e como não se tinham dados de treino para a classe “Queimado”, recorreu-se a uma metodologia para a identificação das zonas queimadas utilizando o NBR. Foi criada uma máscara NBR com as áreas queimadas e foram removidos dos dados de treino os polígonos que estavam intersetados com a classe “Queimado”, de forma a minimizar posteriores erros. Foi multiplicada a classificação dos mapas de cobertura do solo com a máscara NBR. Com esta multiplicação, foi notado que zonas com solo nu à beira rio geraram erros na caracterização dos pixéis. Este problema esteve mais presente na imagem de verão, devido ao nível de água estar mais baixo e a classe de “Solo Nu” ser mais saliente, sendo por isso mais confundida com a classe “Queimado” do que a imagem de outono. Para a resolução deste problema foi utilizada uma ferramenta no ambiente ArcGIS que permitiu a suavização dos resultados. Por fim, a classificação da ocupação do solo foi multiplicada com a máscara NBR, obtendo assim 6 classes:

- Classe 0 – Queimado,
- Classe 1 – Territórios Artificializados,
- Classe 2 – Vegetação herbácea e arbustiva,
- Classe 3 – Floresta,
- Classe 4 – Solo Nu,
- Classe 5 – Água.

Para a obtenção da máscara de florestas, foi simplificada a classificação final em duas classes: na classe “Floresta” e na classe “Não Floresta”. As máscaras de floresta para o inverno e para o verão foram validadas. Para a máscara de floresta do verão a exatidão global foi de 58% e para a máscara de floresta do inverno foi de 63%. Notou-se que os resultados poderiam apresentar

melhores resultados, tanto para o verão como para o inverno, isto deveu-se a três principais razões, aos dados estarem em zonas: perto de caminhos rurais, em plantações jovens e em florestas dispersas. A recolha dos dados nas zonas referidas anteriormente gerou confusão entre classes, entre a classe “Floresta” com as classes “Vegetação herbácea e arbustiva” e “Solo Nu”.

Os erros podem também estar associados ao mau posicionamento dos pontos, isto porque, a presença de árvores interfere com o sinal GPS.

6.2. Identificação automática de alterações nas florestas

A identificação automática de alterações nas florestas passou pelos seguintes passos:

- Obtenção das alterações pela ferramenta disponível do portal F-Tep (método 1),
- Obtenção das alterações pela diferença de máscaras de floresta (método 2),
- Obtenção das alterações pela interseção dos métodos referidos anteriormente (método 3),
- Escolha do melhor método.

A comparação entre métodos, para testar o que apresentava melhor desempenho, deu-se pela análise de imagens de satélite. Nesta análise foram notadas zonas que nos métodos 1 e 2 estavam a ser consideradas como alteração e na realidade não eram. O método 3 minimizou os erros de zonas mal classificadas. Assim sendo, a obtenção das alterações ocorridas na floresta pelo método 3 foi o que mostrou melhor desempenho.

Depois de definido o método com melhor desempenho, este foi validado. Para as alterações de floresta ocorridas entre o outono e o inverno a exatidão global foi de 63%, entre o verão e o inverno foi de 57% e entre o verão e o outono foi de 41%. Os erros existentes deveram-se a pixéis que estavam a ser considerados “Floresta” e na realidade pertenciam à classe “Vegetação herbácea e arbustiva”. Estes erros seriam minimizados com a criação de dados de treino nas zonas erroneamente classificadas, assim estas, deixariam de ser confundidas com a classe “Floresta”.

6.3. Limitações & Trabalho futuro

Devido à falta de tempo, foram utilizados dados de referência já recolhidos pelo RAIZ para a validação das máscaras de floresta e de alterações de floresta. A implementação dos métodos foi adaptada ao tempo limite existente, não entrando em processos de teste demorosos.

Como trabalho futuro:

- 1) Seriam testadas todas as ferramentas do portal F-Tep;
- 2) Seriam criados dados de referência, em vez de se utilizar os fornecidos pelo RAIZ;
- 3) Seriam validados todos os produtos gerados;
- 4) Seriam desenvolvidos mais processos para um conhecimento superior a nível florestal do caso de estudo;
- 5) Em vez de se ter utilizado o ArcGIS, seria testado o QGIS que a plataforma F-Tep possui.

Referências

- 1) Bannari, A., Asalhi, H., & Teillet, P. M. (2002). Transformed difference vegetation index (TDVI) for vegetation cover mapping. *IEEE International Geoscience and Remote Sensing Symposium*, 5, 3053–3055 vol.5. <https://doi.org/10.1109/IGARSS.2002.1026867>
- 2) Cenggoro, T. W., Isa, S. M., & Kusuma, G. P. (2016). Construction of Jakarta Land Use/Land Cover dataset using classification method. *2016 IEEE Region 10 Symposium (TENSYMP), Region 10 Symposium (TENSYMP), 2016 IEEE*, 337–342. <https://doi.org/10.1109/TENCONSpring.2016.7519429>
- 3) *Copernicus Global Land Service*. (sem data). Obtido 15 de Março de 2020, de <https://land.copernicus.eu/global/>
- 4) Delwart, S. (sem data). *ESA Standard Document*. 1, 64.
- 5) Driese, K. (2012). *Vegetation Indices*. 48.
- 6) Escuin, S., Navarro, R., & Fernández, P. (2008). Fire severity assessment by using NBR (Normalized Burn Ratio) and NDVI (Normalized Difference Vegetation Index) derived from LANDSAT TM/ETM images. *International Journal of Remote Sensing*, 29(4), 1053–1073. <https://doi.org/10.1080/01431160701281072>
- 7) *Forest Monitoring, Land Use & Deforestation Trends | Global Forest Watch*. (sem data). Obtido 7 de Março de 2020, de <https://www.globalforestwatch.org/>
- 8) *ForestMonitor*. (sem data). Obtido 30 de Janeiro de 2020, de <https://www.terramonitor.com/products/forestmonitor>
- 9) *Forestry TEP*. (sem data). Obtido 10 de Abril de 2020, de <https://f-tep.com/>
- 10) Garcia, H. M. V. G. dos S. (2017). *A floresta em Portugal. Causas e consequências da expansão do Eucalipto. Caso de estudo: O concelho de Torres Vedras*.
- 11) Gatti, A., & Naud, C. (2017). *Sentinel-2 Products Specification Document*. 487.

- 12) Hedley, J. D., Roelfsema, C., Brando, V., Giardino, C., Kutser, T., Phinn, S., Mumby, P. J., Barrilero, O., Laporte, J., & Koetz, B. (2018). Coral reef applications of Sentinel-2: Coverage, characteristics, bathymetry and benthic mapping with comparison to Landsat 8. *Remote Sensing of Environment*, 216, 598–614. <https://doi.org/10.1016/j.rse.2018.07.014>
- 13) *INPE 8465.pdf*. (sem data). Obtido 25 de Agosto de 2020, de <http://mtc-m12.sid.inpe.br/col/sid.inpe.br/sergio/2004/10.20.14.47/doc/INPE%208465.pdf>
- 14) Meneses, B. M., Reis, E., Vale, M. J., & Reis, R. (2018). Modelling Land Use and Land Cover Changes in Portugal: A Multi-Scale and Multi-Temporal Approach. *LES TYPES D'UTILISATION ET D'OCCUPATION DO SOL AU PORTUGAL. UNE MÉTHODOLOGIE MULTI-SCALAIRE ET MULTI-TEMPORALE.*, 53(107), 3–26. <https://doi.org/10.18055/Finis12258>
- 15) Miettinen, J. (2019). *Forestry TEP User Manual*. 58.
- 16) RAIZ. (2020). *Instituto de investigação da Floresta e do Papel | RAIZ*. Raiz Website. <http://raiz-iifp.pt/>
- 17) Ríos, I., Bouza, P. J., Bortolus, A., & Alvarez, M. del P. (2018). Soil-geomorphology relationships and landscape evolution in a southwestern Atlantic tidal salt marsh in Patagonia, Argentina. *Journal of South American Earth Sciences*, 84, 385–398. <https://doi.org/10.1016/j.jsames.2018.04.015>
- 18) *Sentinel Hub*. (sem data). Obtido 10 de Março de 2020, de <https://www.sentinel-hub.com/>
- 19) Stehman, S. V. (2009). Sampling designs for accuracy assessment of land cover. *International Journal of Remote Sensing*, 30(20), 5243–5272. <https://doi.org/10.1080/01431160903131000>
- 20) *Terramonitor*. (sem data). Obtido 1 de Março de 2020, de <https://www.terramonitor.com/>
- 21) Torres-Moreno, J.-M., Bougrain, L., & Alexandre, F. (2009). Combining Supervised and Unsupervised Learning for GIS Classification. *arXiv:0905.2347 [cs]*. <http://arxiv.org/abs/0905.2347>

- 22) Universidade de Coimbra, EPSSI. (2020, Agosto 21). *Instituto de Engenharia de Sistemas e Computadores de Coimbra*. Instituto de Engenharia de Sistemas e Computadores de Coimbra. /en/org/inescc/Projects/projects/EPSSI
- 23) Hedley, J. D., Roelfsema, C., Brando, V., Giardino, C., Kutser, T., Phinn, S., Mumby, P. J., Barrilero, O., Laporte, J., & Koetz, B. (2018). Coral reef applications of Sentinel-2: Coverage, characteristics, bathymetry and benthic mapping with comparison to Landsat 8. *Remote Sensing of Environment*, 216, 598–614. <https://doi.org/10.1016/j.rse.2018.07.014>