



UNIVERSIDADE D
COIMBRA

Ismael José Fervença de Jesus

**METODOLOGIAS AUTOMÁTICAS PARA CRIAÇÃO DE
DADOS DE TREINO A PARTIR DA
COS 2015, UA 2012 E OSM**

**Dissertação no âmbito do Mestrado em Engenharia de Informação Geoespacial
orientada 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
Tecnologias da Universidade de Coimbra**

Setembro de 2020

Faculdade de Ciências e Tecnologias da Universidade de Coimbra

Metodologias Automáticas para Criação de Dados de Treino a Partir da COS 2015, UA 2012 e OSM

Ismael José Ferverça de Jesus

Dissertação de Mestrado na área científica de Engenharia de Informação Geoespacial orientada pela Doutora Cidália Maria Parreira da Costa Fonte e coorientada pelo 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

Agradecimentos

- ◆ Aos meus orientadores, Professora Doutora Cidália Fonte e Professora Doutor Diogo Duarte, pela sua orientação, pelo tempo investido em mim e pela total colaboração. O apoio e dedicação foram indispensáveis.
- ◆ Ao Joaquim Patriarca, pela disponibilidade, troca de ideias e de saberes.
- ◆ À minha família, por apoiarem e acreditarem em mim nas mudanças que ocorreram ao longo da minha vida académica.
- ◆ Aos meus amigos, especialmente à Carolina, à Joana, à Liliana, à Jéssica, à Raquel, à Beatriz, à Ana Ferreira, à Filipa, à Ana Duarte, à Ana Catarina, ao David, ao Tomás e ao Tiago, pela amizade, companheirismo, força e por todos os momentos passados.
- ◆ Ao INESC Coimbra, pelo financiamento concedido.

Um obrigado a todos!

Resumo

Os mapas de uso e cobertura do solo (MUCS) permitem conhecer o território e geri-lo da melhor forma possível. Com o avanço das tecnologias, pretende-se criar estes mapas de forma rápida, mantendo-os atualizados. Assim, é necessário automatizar todo o seu processo de produção. Uma das fases a automatizar é a criação de dados de treino para a classificação de imagens de satélite, nomeadamente as imagens do Sentinel-2.

Nesta dissertação são abordadas metodologias de criação de dados de treino a partir de dados classificados e disponibilizados de forma gratuita: a Carta e Uso e Ocupação do Solo (COS), o Urban Atlas (UA) e o OpenStreetMap (OSM). Utilizando os dados de treino extraídos a partir destas fontes realizam-se classificações de imagens de satélite da missão Sentinel-2 de 2018 pretendendo assim obter-se MUCS com características semelhantes à COS 2018. Com este objetivo, os dados de treino testados são extraídos exclusivamente a partir da COS 2015, do OSM, do UA 2012 e em simultâneo da COS 2015 e do OSM.

Para os vários dados utilizam-se critérios de decisão que determinam que pixéis são incluídos nas áreas de treino: Presença, Predominância e Exclusividade. Aos dados da COS 2015 são ainda aplicadas filtragens utilizando índices radiométricos. Para se obterem os dados de treino, às áreas de treino são associadas variáveis, *features*, que as caracterizam: as bandas de 10m do Sentinel-2, todas as bandas do Sentinel-2 e ainda todas as bandas do Sentinel-2 e alguns índices radiométricos, sendo analisadas quais as abordagens que produzem melhores resultados quando comparados com a COS 2018.

Das metodologias abordadas, foram obtidos melhores resultados quando os dados de treino são extraídos exclusivamente da COS 2015 não realizando filtragens, tendo como *features* todas as bandas de Sentinel-2.

Palavras-Chave: Dados de treino, Classificação de imagens satélite, Sentinel-2, Mapas de uso e cobertura do solo.

Abstract

The land use land cover maps (LULCM) allow to know the territory and to manage it in the best way. With the advancement of technologies, these maps need to be created quickly, so that they can be kept updated. Therefore, it is necessary to automate the entire production process. One of the phases that need to be automated is the creation of training data for the classification of the satellite images, namely the Sentinel-2 images.

In this work, methodologies will be used for creating training data using classified and freely available data: Carta e Uso e Ocupação do Solo (COS), Urban Atlas (UA) and OpenStreetMap (OSM). With these training data, classifications of Sentinel-2 satellite images from 2018 are carried out, thus aiming to obtain LULCM with characteristics similar to the COS 2018. With this objective, the training data are extracted exclusively from COS 2015, OSM, UA 2012 and simultaneously from COS 2015 and OSM.

For the various data, decision criteria are used to determine which pixels are included in the training areas: Presence, Predominance and Exclusivity. Filters using radiometric indices were also applied to the data of COS 2015. To obtain the training data, the training areas are associated with variables, features, which characterize them: the 10m bands of Sentinel-2, all bands of Sentinel-2 and also all bands of Sentinel-2 and various radiometric indices, and then analyzed the ones that produced better results when compared to COS 2018.

Of the approached methodologies, the best results were obtained with the training data extracted exclusively from COS 2015, with no filtering, having as features all the Sentinel-2 bands.

Keywords: *Training data, Classifications of satellite images, Sentinel-2, Land use land cover maps.*

Índice

Agradecimentos	iii
Resumo.....	v
Abstract	viii
Índice.....	ix
Índice de Figuras.....	xiii
Índice de Tabelas	xxi
Abreviaturas	xxxiii
1. Introdução.....	1
1.1. Motivação e Objetivos	2
1.2. Estrutura	3
2. Produção de Mapas de Uso e Cobertura do Solo	5
2.1. Processo de Criação de um Mapa de Uso e Cobertura do Solo	5
2.2. Missão Sentinel-2.....	7
2.3. Fontes de Dados Classificados.....	8
2.3.1. Carta de Ocupação e Uso do Solo.....	9
2.3.2. Urban Atlas	11
2.3.3. Informação Geográfica Voluntária - OpenStreetMap.....	13
2.4. Classificadores de Dados	15
2.5. Índices Radiométricos.....	18
2.5.1. NDVI.....	19
2.5.2. NDWI.....	20
2.5.3. NDBI.....	20

2.5.4. SAVI.....	21
2.6. Validação de MUCSs	21
3. Metodologias de Criação Automática de Dados de Treino.....	25
3.1. Definição e Harmonização da Nomenclatura.....	26
3.2. Criação de Grelhas	30
3.3. Criação de Dados de Treino	32
3.3.1. Criação de Áreas de Treino.....	32
3.3.2. Associação de Features às Áreas de Treino	34
3.3.3. Processo de Filtragem das Áreas de Treino com Índices Radiométricos	35
3.3.4. Denominação das Áreas de Treino e Dados de Treino	36
3.4. Avaliação dos Dados de Treino	37
3.5. Classificação e Generalização.....	38
3.6. Validação.....	39
4. Casos de Estudo	41
4.1. Áreas de Estudo.....	41
4.1.1. Área de Estudo A: Lisboa	41
4.1.2. Área de Estudo B: Serra da Estrela	44
4.2. Resultados e Discussão	48
4.2.1. Classificação com dados de treino extraídos da COS 2015	48
4.2.2. Classificação com dados de treino extraídos da COS 2015 com filtragens a partir dos índices radiométricos.....	61
4.2.3. Classificação com dados de treino extraídos do OSM.....	69
4.2.4. Classificação com dados de treino extraídos da COS 2015 e do OSM.....	77
4.2.5. Classificação com dados de treino extraídos do UA 2012.....	84
4.2.6. Classificação com dados de treino extraídos da COS 2015 Utilizando os Vizinhos como Features	88
Conclusão	93
Bibliografia.....	97
Anexos.....	103

Anexo A – Criação de Grelhas	105
Anexo B – Preenchimento das Grelhas	107
Anexo C – Criação das Áreas de Treino a Partir da COS 2015	111
Anexo D – Criação das Amostras das Áreas de Treino a Partir da COS 2015	117
Anexo E – MUCS de Referência.....	121
Anexo F – Classificações e Respetivas Generalizações a Partir dos Dados de Treino Extraídos da COS 2015.....	123
Anexo G – Histogramas dos Índices NDVI, NDWI e NDBI por Classe	143
Anexo H – Classificações e Respetivas Generalizações a Partir dos Dados de Treino Obtidos pela Filtragem dos Dados de Treino Extraídos da COS 2015	153
Anexo I – Classificações e Respetivas Generalizações a Partir dos Dados de Treino Extraídos do OSM.....	159
Anexo J – Classificações e respetivas generalizações a partir dos dados de treino obtidos em simultâneo pela COS 2015 e pelo OSM.....	167
Anexo K – Classificações e respetivas generalizações a partir dos dados de treino extraídos do UA 2012.....	175
Anexo L – Classificações e Respetivas Generalizações a Partir dos Dados de Treino Extraídos da COS 2015 Utilizando os Vizinhos como Features	179

Índice de Figuras

Figura 1: Nível 1 das COS 2015 e COS 2018 para Portugal Continental	9
Figura 2: Cidades abrangidas pelo UA 2012 para Portugal Continental e para parte de Espanha (Fonte: https://land.copernicus.eu/local/urban-atlas/urban-atlas-2012).....	12
Figura 3: Curvas de refletância da água límpida, água turva, da vegetação e de solo seco (adaptado de Dalezios, 2014).....	16
Figura 4: Exemplo de árvore de decisão. (Foto cedida por Tiago Rocha).....	17
Figura 5: Etapas principais da metodologia e secções do texto onde são apresentadas.	26
Figura 6: Fluxograma que ilustra o processo de criação das grelhas.....	31
Figura 7: Fluxograma que ilustra o processo de criação de dados de treino (sem filtragem com índices radiométricos).	34
Figura 8: Fluxograma que ilustra o processo de criação de áreas de treino com filtragem usando índices radiométricos.....	36
Figura 9: Fluxograma que ilustra o processo de classificação e sua generalização.....	39
Figura 10: Fluxograma que ilustra o processo de validação dos MUCSs criados.	40
Figura 11: Área de estudo A – a) enquadramento da área de estudo em Portugal Continental (Basemap: Bing Aerial); composições a cor natural, RGB (bandas 4-3-2), das imagens de Sentinel-2 para as datas referentes aos meses de março (b), de junho (c) e outubro (d) de 2018.	42
Figura 12: Área de estudo A – Nível 1 da COS 2015.....	43
Figura 13: Área de estudo A – Nível 1 da COS 2018.....	43
Figura 14: Área de estudo A – OpenStreetMap (Basemap OpenStreetMap no QGIS).....	44
Figura 15: Área de estudo A – Nível 1 do UA 2012.....	44
Figura 16: Área de estudo B – a) enquadramento da área de estudo em Portugal Continental (Basemap: Bing Aerial); composições a cor natural, RGB (bandas 4-3-2), das imagens de Sentinel-2 para as datas: b) Março de 2018; c) Junho de 2018; d) Outubro de 2018.	45
Figura 17: Área de estudo B – Nível 1 da COS 2015.....	46
Figura 18: Área de estudo B – Nível 1 da COS 2018.....	47
Figura 19: Área de estudo B – OpenStreetMap (Basemap QGIS: OpenStreetMap).....	47
Figura 20: Área de estudo A – Áreas cobertas pelos dados de treino obtidos a partir da COS 2015 (azul – área coberta por DT_C0 e DT_C1; amarelo – área pertencente a DT_C0, DT_C1 e DT_C2).....	49

Figura 21: Área de estudo B – Áreas cobertas pelos dados de treino obtidos a partir da COS 2015 (azul – área coberta por DT_C0 e DT_C1; amarelo – área pertencente a DT_C0, DT_C1 e DT_C2).....	49
Figura 22: Área de estudo A – Avaliação da separabilidade da amostra de dados de treino obtidos diretamente pela COS 2015 usando a distância de Bhattacharya.....	52
Figura 23: Área de estudo B – Avaliação da separabilidade da amostra de dados de treino obtidos diretamente pela COS 2015 usando a distância de Bhattacharya.....	52
Figura 24: Região escolhida ao acaso na área de estudo A – MUCSs obtidos pelas classificações e respectivas generalizações.....	55
Figura 25: Região escolhida ao acaso na área de estudo B – MUCSs obtidos pelas classificações e respectivas generalizações.....	56
Figura 26: Área de estudo A: Mudança de uso e ocupação do solo entre a COS 2015 e a COS 2018.	60
Figura 27: Área de estudo B: Mudança de uso e ocupação do solo entre a COS 2015 e a COS 2018.	60
Figura 28: Área de estudo A – Avaliação da separabilidade da amostra de dados de treino obtidos através da COS 2015 com filtragens a partir dos índices radiométricos usando a distância de Bhattacharya.	64
Figura 29: Área de estudo B – Avaliação da separabilidade da amostra de dados de treino obtidos através da COS 2015 com filtragens a partir dos índices radiométricos usando a distância de Bhattacharya.	64
Figura 30: Região escolhida ao acaso na área de estudo A – MUCSs obtidos pelas classificações e respectivas generalizações a partir dos dados de treino criados com filtragens usando os índices radiométricos.	65
Figura 31: Região escolhida ao acaso na área de estudo B – MUCSs obtidos pelas classificações e respectivas generalizações a partir dos dados de treino criados com filtragens usando os índices radiométricos.	66
Figura 32: Área de estudo A – Áreas cobertas pelos dados de treino obtidos a partir do OSM (azul – área coberta por DT_O0 e DT_O1; amarelo – área pertencente a DT_O0, DT_O1 e DT_O2).	70
Figura 33: Área de estudo B – Áreas cobertas pelos dados de treino obtidos a partir do OSM (azul – área coberta por DT_O0 e DT_O1; amarelo – área pertencente a DT_O0, DT_O1 e DT_O2).	70
Figura 34: Região escolhida ao acaso na área de estudo A – MUCSs obtidos pelas classificações e respectivas generalizações a partir dos dados de treino criados a partir do OSM.	73
Figura 35: Região escolhida ao acaso na área de estudo B – MUCSs obtidos pelas classificações e respectivas generalizações a partir dos dados de treino criados a partir do OSM.	74

Figura 36: Região escolhida ao acaso na área de estudo A – MUCSs obtidos pelas classificações e respectivas generalizações a partir dos dados de treino criados a partir da COS 2015 e do OSM.	80
Figura 37: Região escolhida ao acaso na área de estudo B – MUCSs obtidos pelas classificações e respectivas generalizações a partir dos dados de treino criados a partir da COS 2015 e do OSM.	81
Figura 38: Área de estudo A – Áreas cobertas pelos dados de treino obtidos a partir do UA 2012 (azul – área coberta por DT_U0 e DT_U1; amarelo – área pertencente a DT_U0, DT_U1 e DT_U2).	84
Figura 39: Região escolhida ao acaso na área de estudo A – MUCSs obtidos pelas classificações e respectivas generalizações a partir dos dados de treino criados a partir do UA 2012.	86
Figura 40: Região escolhida ao acaso na área de estudo A – MUCSs obtidos pelas classificações a partir dos dados de treino criados a partir do COS 2015 utilizando como <i>features</i> as bandas de 10m no pixel e nos pixels vizinhos (Norte, Sul, Este e Oeste).....	89
Figura 41: Região escolhida ao acaso na área de estudo B – MUCSs obtidos pelas classificações a partir dos dados de treino criados a partir do COS 2015 utilizando como <i>features</i> as bandas de 10m no pixel e nos pixels vizinhos (Norte, Sul, Este e Oeste).....	89
Figura 42: Área de Estudo A – Mapa de Referência 1.	121
Figura 43: Área de Estudo A – Mapa de Referência 2.	122
Figura 44: Área de Estudo B – Mapa de Referência.....	122
Figura 45: Área de Estudo A – MUCS obtido pela classificação com DT_C0_F1.....	124
Figura 46: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C0_F1.	124
Figura 47: Área de Estudo A – MUCS obtido pela classificação com DT_C0_F2.....	125
Figura 48: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C0_F2.	125
Figura 49: Área de Estudo A – MUCS obtido pela classificação com DT_C0_F3.....	126
Figura 50: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C0_F3.	126
Figura 51: Área de Estudo A – MUCS obtido pela classificação com DT_C1_F1.....	127
Figura 52: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C1_F1.	127
Figura 53: Área de Estudo A – MUCS obtido pela classificação com DT_C1_F2.....	128
Figura 54: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C1_F2.	128
Figura 55: Área de Estudo A – MUCS obtido pela classificação com DT_C1_F3.....	129

Figura 56: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C1_F3.	129
Figura 57: Área de Estudo A – MUCS obtido pela classificação com DT_C2_F1.	130
Figura 58: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C2_F1.	130
Figura 59: Área de Estudo A – MUCS obtido pela classificação com DT_C2_F2.	131
Figura 60: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C2_F2.	131
Figura 61: Área de Estudo A – MUCS obtido pela classificação com DT_C2_F3.	132
Figura 62: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C2_F3.	132
Figura 63: Área de Estudo B – MUCS obtido pela classificação com DT_C0_F1.	133
Figura 64: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C0_F1.	133
Figura 65: Área de Estudo B – MUCS obtido pela classificação com DT_C0_F2.	134
Figura 66: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C0_F2.	134
Figura 67: Área de Estudo B – MUCS obtido pela classificação com DT_C0_F3.	135
Figura 68: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C0_F3.	135
Figura 69: Área de Estudo B – MUCS obtido pela classificação com DT_C1_F1.	136
Figura 70: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C1_F1.	136
Figura 71: Área de Estudo B – MUCS obtido pela classificação com DT_C1_F2.	137
Figura 72: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C1_F2.	137
Figura 73: Área de Estudo B – MUCS obtido pela classificação com DT_C1_F3.	138
Figura 74: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C1_F3.	138
Figura 75: Área de Estudo B – MUCS obtido pela classificação com DT_C2_F1.	139
Figura 76: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C2_F1.	139
Figura 77: Área de Estudo B – MUCS obtido pela classificação com DT_C2_F2.	140
Figura 78: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C2_F2.	140
Figura 79: Área de Estudo B – MUCS obtido pela classificação com DT_C2_F3.	141

Figura 80: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C2_F3.	141
Figura 81: Área de Estudo A – Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 1.....	144
Figura 82: Área de Estudo B - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 1.....	144
Figura 83: Área de Estudo A - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 2.....	145
Figura 84: Área de Estudo B - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 2.....	145
Figura 85: Área de Estudo A - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 3.....	146
Figura 86: Área de Estudo B - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 3.....	146
Figura 87: Área de Estudo A - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 4.....	147
Figura 88: Área de Estudo B - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 4.....	147
Figura 89: Área de Estudo A - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 5.....	148
Figura 90: Área de Estudo B - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 5.....	148
Figura 91: Área de Estudo A - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 6.....	149
Figura 92: Área de Estudo B - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 6.....	149
Figura 93: Área de Estudo A - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 7.....	150
Figura 94: Área de Estudo A - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 8.....	150
Figura 95: Área de Estudo B - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 8.....	151
Figura 96: Área de Estudo A – MUCS obtido pela classificação com DT_C1L_F2.	154
Figura 97: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C1L_F2.	154
Figura 98: Área de Estudo A – MUCS obtido pela classificação com DT_C1M_F2.	155

Figura 99: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C1M_F2.	155
Figura 100: Área de Estudo B – MUCS obtido pela classificação com DT_C1L_F2.	156
Figura 101: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C1L_F2.	156
Figura 102: Área de Estudo B – MUCS obtido pela classificação com DT_C1M_F2.	157
Figura 103: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C1M_F2.	157
Figura 104: Área de Estudo A – MUCS obtido pela classificação com DT_O0_F2.	160
Figura 105: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_O0_F2.	160
Figura 106: Área de Estudo A – MUCS obtido pela classificação com DT_O1_F2.	161
Figura 107: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_O1_F2.	161
Figura 108: Área de Estudo A – MUCS obtido pela classificação com DT_O2_F2.	162
Figura 109: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_O2_F2.	162
Figura 110: Área de Estudo B – MUCS obtido pela classificação com DT_O0_F2.	163
Figura 111: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_O0_F2.	163
Figura 112: Área de Estudo B – MUCS obtido pela classificação com DT_O1_F2.	164
Figura 113: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_O1_F2.	164
Figura 114: Área de Estudo B – MUCS obtido pela classificação com DT_O2_F2.	165
Figura 115: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_O2_F2.	165
Figura 116: Área de Estudo A – MUCS obtido pela classificação com DT_CO0_F2.	168
Figura 117: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_CO0_F2.	168
Figura 118: Área de Estudo A – MUCS obtido pela classificação com DT_CO1_F2.	169
Figura 119: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_CO1_F2.	169
Figura 120: Área de Estudo A – MUCS obtido pela classificação com DT_CO2_F2.	170
Figura 121: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_CO2_F2.	170
Figura 122: Área de Estudo B – MUCS obtido pela classificação com DT_CO0_F2.	171

Figura 123: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_CO0_F2.....	171
Figura 124: Área de Estudo B – MUCS obtido pela classificação com DT_CO1_F2.	172
Figura 125: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_CO1_F2.....	172
Figura 126: Área de Estudo B – MUCS obtido pela classificação com DT_CO2_F2.	173
Figura 127: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_CO2_F2.....	173
Figura 128: Área de Estudo A – MUCS obtido pela classificação com DT_U0_F2.....	176
Figura 129: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_U0_F2.	176
Figura 130: Área de Estudo A – MUCS obtido pela classificação com DT_U1_F2.....	177
Figura 131: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_U1_F2.	177
Figura 132: Área de Estudo A – MUCS obtido pela classificação com DT_U2_F2.....	178
Figura 133: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_U2_F2.	178
Figura 134: Área de Estudo A – MUCS obtido pela classificação com DT_C1_F1-4.	180
Figura 135: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C1_F1-4.....	180
Figura 136: Área de Estudo B – MUCS obtido pela classificação com DT_C1_F1-4.....	181
Figura 137: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C1_F1-4.....	181

Índice de Tabelas

Tabela 1: Bandas multiespectrais do Sentinel-2.	7
Tabela 2: Matriz de confusão para um problema de n classes.	23
Tabela 3: Harmonização entre a nomenclatura a adotar e as classes da COS 2015 e as classes da COS 2018.	28
Tabela 4: Harmonização entre a nomenclatura a adotar e as classes do UA 2012 e as classes da COS 2018. Observação: a classe 5 deixa de existir e é agregada à classe 3.	29
Tabela 5: Harmonização entre a nomenclatura a adotar e a nomenclatura do CLC adotada no software OSM2LULC.	30
Tabela 6: Denominações dos dados de treino consoante os critérios e <i>features</i> utilizados.	37
Tabela 7: Detalhes das imagens de Sentinel-2 utilizadas, para a área de estudo A.	42
Tabela 8: Detalhes das imagens de Sentinel-2 utilizadas para a área de estudo B.	45
Tabela 9: Área de estudo A – Validação dos dados de treino criados a partir da COS 2015 e respetivas amostras: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos para DT_C0, DT_C1 e DT_C2, e para as amostras deles extraídas, denotadas, respetivamente, por S_DT_C0, S_DT_C1 e S_DT_C2.	50
Tabela 10: Área de estudo B – Validação dos dados de treino criados a partir da COS 2015 e respetivas amostras: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos para DT_C0, DT_C1 e DT_C2, e para as amostras deles extraídas, denotadas, respetivamente, por S_DT_C0, S_DT_C1 e S_DT_C2.	51
Tabela 11: Área de Estudo A – Exatidão global das classificações utilizando os dados de treino obtidos a partir da COS 2015.	53
Tabela 12: Área de Estudo A – Exatidão global das generalizações das classificações utilizando os dados de treino obtidos a partir da COS 2015.	53
Tabela 13: Área de Estudo B – Exatidão global das classificações utilizando os dados de treino obtidos a partir da COS 2015.	54
Tabela 14: Área de Estudo B – Exatidão global das generalizações das classificações utilizando os dados de treino obtidos a partir da COS 2015.	54
Tabela 15: Área de Estudo A: Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_C0_F2, DT_C1_F2 e DT_C2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas	

decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)	57
Tabela 16: Área de Estudo A: Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_C0_F2, DT_C1_F2 e DT_C2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)	58
Tabela 17: Área de Estudo B: Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_C0_F2, DT_C1_F2 e DT_C2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)	59
Tabela 18: Área de Estudo B: Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_C0_F2, DT_C1_F2 e DT_C2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)	59
Tabela 19: Área de estudo A: Percentagem de área corretamente classificada nas áreas onde ocorreram mudança de ocupação e uso do solo, da COS 2015 para COS 2018.	61
Tabela 20: Área de estudo B: Percentagem de área corretamente classificada nas áreas onde ocorreram mudança de ocupação e uso do solo, da COS 2015 para COS 2018.	61
Tabela 21: Condições a verificar por cada classe a partir dos índices radiométricos NDVI, NDWI e NDBI.	62
Tabela 22: Área de estudo A – Validação das amostras dos dados de treino criados a partir da filtragem com os índices radiométricos das áreas de treino DT_C1: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos. Denota-se por S_DT_C1, S_DT_C1L e S_DT_C1M as amostras de DT_C1, DT_C1L e DT_C1M, respetivamente.	63
Tabela 23: Área de estudo B – Validação das amostras dos dados de treino criados a partir da filtragem com os índices radiométricos das áreas de treino DT_C1: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos. Denota-se por S_DT_C1, S_DT_C1L e S_DT_C1M as amostras de DT_C1, DT_C1L e DT_C1M, respetivamente.	63
Tabela 24: Exatidões globais das classificações, e respetivas generalizações, utilizando os dados de treino DT_C1_F2 com filtragens utilizando os índices radiométricos com os critérios L e M.	67

Tabela 25: Área de Estudo A – Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_C1_F2, DT_C1L_F2 e DT_C1M_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)	67
Tabela 26: Área de Estudo A – Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_C1_F2, DT_C1L_F2 e DT_C1M_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)	68
Tabela 27: Área de Estudo B – Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_C1_F2, DT_C1L_F2 e DT_C1M_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)	68
Tabela 28: Área de Estudo B – Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_C1_F2, DT_C1L_F2 e DT_C1M_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)	69
Tabela 29: Área de estudo A – Validação das amostras dos dados de treino criados a partir OSM: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos. Denota-se por S_DT_O0, S_DT_O1 e S_DT_O2 as amostras de DT_O0, DT_O1 e DT_O2, respetivamente.....	71
Tabela 30: Área de estudo B – Validação das amostras dos dados de treino criados a partir OSM: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos. Denota-se por S_DT_O0, S_DT_O1 e S_DT_O2 as amostras de DT_O0, DT_O1 e DT_O2, respetivamente.....	71
Tabela 31: Exatidões globais das classificações, e respetivas generalizações, utilizando os dados de treino DT_O0_F2, DT_O1_F2 e DT_O2_F2.	72
Tabela 32: Área de Estudo A – Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_O0_F2, DT_O1_F2 e DT_O2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-	

se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)	75
Tabela 33: Área de Estudo A – Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_O0_F2, DT_O1_F2 e DT_O2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)	75
Tabela 34: Área de Estudo B – Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_O0_F2, DT_O1_F2 e DT_O2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)	76
Tabela 35: Área de Estudo B – Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_O0_F2, DT_O1_F2 e DT_O2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)	76
Tabela 36: Área de estudo A – Validação das amostras dos dados de treino criados em simultâneo a partir OSM e da COS 2015: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos.	77
Tabela 37: Área de estudo B – Validação das amostras dos dados de treino criados em simultâneo a partir OSM e da COS 2015: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos.	77
Tabela 38: Área de estudo A – Validação das amostras dos dados de treino criados a partir OSM, da COS 2015 e em simultâneo por ambas as fontes de dados: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos. Para facilitar a leitura, os valores relativos ao uso em simultâneo da COS 2015 e do OSM encontram-se sombreados a verde quando são melhores que os valores relativos aos dados de treino criados exclusivamente a partir de cada fonte. Caso contrário, estes valores apresentam-se sombreados a vermelho.	78
Tabela 39: Área de estudo B – Validação das amostras dos dados de treino criados a partir OSM, da COS 2015 e em simultâneo por ambas as fontes de dados: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos. Para facilitar a leitura, os	

valores relativos ao uso em simultâneo da COS 2015 e do OSM encontram-se sombreados a verde quando são melhores que os valores relativos aos dados de treino criados exclusivamente a partir de cada fonte. Caso contrário, estes valores apresentam-se sombreados a vermelho.....	78
Tabela 40: Exatidões globais das classificações, e respetivas generalizações, utilizando os dados de treino DT_CO0_F2, DT_CO1_F2 e DT_CO2_F2.	79
Tabela 41: Área de Estudo A – Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_CO0_F2, DT_CO1_F2 e DT_CO2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização).....	82
Tabela 42: Área de Estudo A – Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_CO0_F2, DT_CO1_F2 e DT_CO2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)	82
Tabela 43: Área de Estudo B – Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_CO0_F2, DT_CO1_F2 e DT_CO2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização).....	83
Tabela 44: Área de Estudo B – Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_CO0_F2, DT_CO1_F2 e DT_CO2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)	83
Tabela 45: Área de estudo A – Validação das amostras dos dados de treino criados a partir do UA 2012: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos. Denota-se por S_DT_U0, S_DT_U1 e S_DT_U2 as amostras de DT_U0, DT_U1 e DT_U2, respetivamente.....	85
Tabela 46: Exatidões globais das classificações, e respetivas generalizações, utilizando os dados de treino DT_U0_F2, DT_U1_F2 e DT_U2_F2.	87

Tabela 47: Área de Estudo A – Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_U0_F2, DT_U1_F2 e DT_U2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)	87
Tabela 48: Área de Estudo A – Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_U0_F2, DT_U1_F2 e DT_U2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)	88
Tabela 49: Exatidões globais das classificações e generalizações utilizando os dados de treino DT_C1_F1 e DT_C1_F1-4.....	90
Tabela 50: Área de Estudo A – Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_C1_F1 e DT_C1_F1-4. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)	90
Tabela 51: Área de Estudo A – Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_C1_F1 e DT_C1_F1-4. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)	91
Tabela 52: Área de Estudo B – Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_C1_F1 e DT_C1_F1-4. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)	91
Tabela 53: Área de Estudo B – Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_C1_F1 e DT_C1_F1-4. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)	92
Tabela 54: Percentagem de área corretamente classificada nas áreas onde ocorreram mudança de ocupação e uso do solo, da COS 2015 para COS 2018.....	92

Tabela 55: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C0_F1.	124
Tabela 56: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C0_F1.....	124
Tabela 57: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C0_F2.	125
Tabela 58: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C0_F2.....	125
Tabela 59: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C0_F3.	126
Tabela 60: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C0_F3.....	126
Tabela 61: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1_F1.	127
Tabela 62: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1_F1.....	127
Tabela 63: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1_F2.	128
Tabela 64: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1_F2.....	128
Tabela 65: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1_F3.	129
Tabela 66: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1_F3.....	129
Tabela 67: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C2_F1.	130
Tabela 68: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C2_F1.....	130
Tabela 69: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C2_F2.	131
Tabela 70: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C2_F2.....	131
Tabela 71: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C2_F3.	132
Tabela 72: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C2_F3.....	132

Tabela 73: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C0_F1.....	133
Tabela 74: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C0_F1.....	133
Tabela 75: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C0_F2.....	134
Tabela 76: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C0_F2.....	134
Tabela 77: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C0_F3.....	135
Tabela 78: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C0_F3.....	135
Tabela 79: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1_F1.....	136
Tabela 80: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1_F1.....	136
Tabela 81: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1_F2.....	137
Tabela 82: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1_F2.....	137
Tabela 83: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1_F3.....	138
Tabela 84: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1_F3.....	138
Tabela 85: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C2_F1.....	139
Tabela 86: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C2_F1.....	139
Tabela 87: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C2_F2.....	140
Tabela 88: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C2_F2.....	140
Tabela 89: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C2_F3.....	141
Tabela 90: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C2_F3.....	141

Tabela 91: Área de Estudo A – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 1.	144
Tabela 92: Área de Estudo B – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 1.	144
Tabela 93: Área de Estudo A – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 2.	145
Tabela 94: Área de Estudo B – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 2.	145
Tabela 95: Área de Estudo A – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 3.	146
Tabela 96: Área de Estudo B – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 3.	146
Tabela 97: Área de Estudo A – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 4.	147
Tabela 98: Área de Estudo B – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 4.	147
Tabela 99: Área de Estudo A – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 5.	148
Tabela 100: Área de Estudo B – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 5.	148
Tabela 101: Área de Estudo A – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 6.	149
Tabela 102: Área de Estudo B – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 6.	149
Tabela 103: Área de Estudo A – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 7.	150
Tabela 104: Área de Estudo A – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 8.	150
Tabela 105: Área de Estudo B – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 8.	151
Tabela 106: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1L_F2.	154
Tabela 107: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1L_F2.	154
Tabela 108: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1M_F2.	155

Tabela 109: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1M_F2.....	155
Tabela 110: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1L_F2.	156
Tabela 111: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1L_F2.....	156
Tabela 112: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1M_F2.	157
Tabela 113: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1M_F2.....	157
Tabela 114: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_O0_F2.....	160
Tabela 115: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_O0_F2.....	160
Tabela 116: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_O1_F2.....	161
Tabela 117: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_O1_F2.....	161
Tabela 118: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_O2_F2.....	162
Tabela 119: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_O2_F2.....	162
Tabela 120: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_O0_F2.....	163
Tabela 121: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_O0_F2.....	163
Tabela 122: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_O1_F2.....	164
Tabela 123: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_O1_F2.....	164
Tabela 124: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_O2_F2.....	165
Tabela 125: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_O2_F2.....	165
Tabela 126: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_CO0_F2.....	168

Tabela 127: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_CO0_F2.	168
Tabela 128: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_CO1_F2.	169
Tabela 129: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_CO1_F2.	169
Tabela 130: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_CO2_F2.	170
Tabela 131: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_CO2_F2.	170
Tabela 132: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_CO0_F2.	171
Tabela 133: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_CO0_F2.	171
Tabela 134: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_CO1_F2.	172
Tabela 135: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_CO1_F2.	172
Tabela 136: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_CO2_F2.	173
Tabela 137: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_CO2_F2.	173
Tabela 138: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_U0_F2.	176
Tabela 139: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_U0_F2.	176
Tabela 140: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_U1_F2.	177
Tabela 141: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_U1_F2.	177
Tabela 142: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_U2_F2.	178
Tabela 143: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_U2_F2.	178
Tabela 144: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1_F1-4.	180

Tabela 145: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1_F1-4.....	180
Tabela 146: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1_F1-4.....	181
Tabela 147: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1_F1-4.....	181

Abreviaturas

COS – Carta de Ocupação e Uso do Solo
CLC – Corine Land Cover
DGT – Direção Geral do Território
IGV – Informação Geográfica Voluntária
MUCS – Mapa de Uso e Cobertura do Solo
OSM – OpenStreet Map
S2 – Sentinel-2
UA – Urban Atlas
UMC – Unidade Mínima Cartográfica

I. Introdução

A necessidade do conhecimento de como o solo da superfície terrestre é ocupado e usado é muito importante para vários tipos de aplicações. Pode-se definir como solo, o conjunto das componentes físicas e biológicas da camada da superfície terrestre. O uso corresponde à utilização que o ser humano dá ao solo, como a utilização de terrenos para agricultura ou para a construção de edifícios. Por outro lado, a cobertura corresponde a atributos físicos da superfície terrestre, como a existência de vegetação ou de água (Regmi et al., 1995). Dada a rápida dinâmica de transformação do uso e ocupação do solo, ao longo do tempo este conhecimento tem tido um papel fundamental nas decisões de como o ser humano deve gerir o solo, tendo um grande impacto nas mais variadíssimas áreas de estudo, como: alterações climáticas (Li et al., 2017; Ren et al., 2015); erosão do solo (Kidane et al., 2019); ordenamento do território (Hersperger et al., 2018); hidrologia (Ahiablame et al., 2017; A. Rajib & Merwade, 2017; M. A. Rajib et al., 2016); epidemiologia (Dlamini et al., 2019) entre muitas outras. Com base nestes estudos, são tomadas decisões que fomentam a sustentabilidade do desenvolvimento sócio-económico, uma vez que as alterações do uso do solo estão diretamente associadas às alterações climáticas, aos processos nos ecossistemas, aos ciclos biogeoquímicos, à biodiversidade e às atividades humanas (Lambin et al., 1999).

Uma forma prática de apresentar a informação do uso e cobertura da superfície terrestre, é com mapas de uso e cobertura do solo (MUCS). Pode-se definir um MUCS como uma representação visual da classificação de uma região quanto à cobertura e uso que o ser humano lhe dá. Com o crescimento da população, também as alterações ao solo têm sido crescentes e rápidas. Assim, é necessário a atualização frequente dos MUCSs para que seja possível tomar decisões baseadas em informações corretas (Bédard et al., 2015; Guptill & Morrison, 1995), o que requer também a validação deste mapas. A avaliação da qualidade temática pretende avaliar o quão próxima é a classificação apresentada nos MUCSs da realidade (Stehman & Foody, 2019).

Um MUCS pode ser criado de forma automática ou semiautomática a partir da classificação de imagens de satélites que recolhem dados com grande frequência sobre a superfície terrestre. Se este processo for feito manualmente a criação de MUCSs torna-se demorada e dispendiosa, não sendo compatível com a existência de mapas atualizados com grande frequência. Com a evolução tecnológica, estes satélites têm sido melhorados e a sua quantidade tem aumentado ao longo do tempo. Desta forma, atualmente existe uma grande quantidade de satélites pertencentes a programas de observação da Terra, nomeadamente o programa Sentinel (Pabian & Institute for Transuranium

Elements, 2015). Acompanhando esta evolução tem-se também a contínua melhoria do poder de computação e de algoritmos de classificação que têm ajudado na produção de MUCS.

I.1. Motivação e Objetivos

Com o avanço da tecnologia, tem-se vindo a recolher dados de observação da Terra com maior resolução temporal. Desta forma, é ambicioso que consigamos criar MUCSs que acompanhem este ritmo de obtenção de imagens. Para que se consiga fazer este acompanhamento, é necessário desenvolver técnicas que automatizem o processo de criação dos MUCS. Quando são usados classificadores supervisionados estes necessitam de um conhecimento prévio de regiões onde são conhecidas as classes a usar na classificação. Estes dados, que têm de ser conhecidos previamente, são denominados dados de treino. Habitualmente, a criação destes dados de treino costuma ser feita manualmente, envolvendo um operador a olhar para imagens e a definir e classificar áreas de treino para cada classe. Esta etapa do processo de criação de MUCSs torna-se lenta e dispendiosa, podendo mesmo envolver visitas de campo em casos de dúvidas. Se esta etapa for feita exclusivamente por fotointerpretação, o processo torna-se subjetivo e influenciado pela acuidade visual e até pelo estado psicológico do utilizador.

Atualmente, está disponível uma grande quantidade de informação sobre a classificação da superfície terrestre usando várias classes, fontes de dados e metodologias. Uma das questões que se pode colocar é se estes dados podem servir como suporte no processo de automatização na criação de dados de treino. Assim, com estes dados de treino seria possível classificar imagens de satélite e obter MUCS. Desta forma, é expectável que com a criação automática dos dados de treino de forma rápida e menos subjetiva se poupe tempo no processo de criação de MUCS, podendo assim obter-se dados de treino, e consequentemente MUCSs atualizados de forma rápida.

O presente trabalho foi feito no âmbito do projeto EPSSI (*Exploring the Potential of the Sentinel missions Satellite Imagery*) financiado pelo INESC Coimbra que tem como objetivo a exploração de imagens de satélite recolhidas pelas missões Sentinel. A realização deste trabalho tem assim como objetivos:

- identificar possíveis fontes de dados que possam servir de base para a criação de dados de treino;
- desenvolver metodologias de filtragem a partir dos dados referidos no ponto anterior de forma a obter dados de treino com qualidade suficiente para a geração de MUCS;

- a partir dos dados de treino gerados automaticamente, classificar imagens de satélite Sentinel-2 (S2) de forma a criar MUCSs que se aproximem da Carta de Ocupação e Uso do Solo (COS) de 2018;
- avaliar a qualidade dos mapas obtidos;
- identificar a(s) metodologia(s) mais promissora(s) para a criação automática de MUCSs.

1.2. Estrutura

A presente dissertação encontra-se assim dividida em quatro capítulos, onde no presente capítulo introduz-se o conceito de MUCS e apresentam-se os objetivos do trabalho desenvolvido. No segundo capítulo explica-se o processo de produção de MUCS, caracterizando produtos de informação geográfica existentes, o programa S2 e os processos de classificação e validação de MUCS. No terceiro capítulo apresentam-se as metodologias usadas para a criação de dados de treino extraídos dos produtos de informação geográfica apresentados no capítulo anterior. Por fim, no quarto capítulo, apresentam-se os resultados de dois casos de estudo referentes a duas áreas distintas de Portugal.

2. Produção de Mapas de Uso e Cobertura do Solo

Um MUCS corresponde a uma representação da classificação de uma região quanto à cobertura e uso que o ser humano lhe dá. Pode-se definir um MUCS como uma aplicação matemática, havendo uma correspondência de cada ponto da região considerada a uma ou mais classes de uso e cobertura de solo, por exemplo, atribuindo uma percentagem de cobertura a cada classe. Os MUCSs mais comuns são constituídos por polígonos (áreas) com correspondência a uma e uma só classe de uso e cobertura/ocupação de solo. Nesta forma, assume-se que a correspondência é homogênea, onde cada ponto do interior do polígono pertence à respetiva classe do polígono.

Um mapa que caracterize o uso e a cobertura que o ser humano dá à superfície terrestre, pode ser criado de forma manual, apoiando-se em fotointerpretação de imagens de satélite e/ou visitas de campo, ou então através da classificação automática destas imagens. No caso dos MUCSs criados a partir da classificação de imagens de satélite, cada pixel da imagem corresponde a uma área no solo a que se faz corresponder uma classe, usando classificadores rígidos (Hu et al., 2013). No caso de se usar classificadores não rígidos, é atribuído a cada pixel uma percentagem, probabilidade ou um grau de pertença a cada uma das classes (Hu et al., 2013).

2.1. Processo de Criação de um Mapa de Uso e Cobertura do Solo

Quando a produção de MUCSs é feita a partir de classificação automática de imagens satélite, usando classificadores supervisionados, podem-se definir as seguintes etapas essenciais do processo:

- definição da finalidade da criação do MUCS;
- definição das classes a adotar;
- seleção da área de estudo e do período de tempo para o qual o MUCS é válido;
- seleção das imagens de satélite que englobem a área em estudo e o período pretendidos;

- análise dos dados a considerar como *features* no processo de classificação, ou seja, análise dos dados que podem caracterizar o terreno para o processo de classificação;
- criação dos dados de treino;
- classificação das imagens;
- avaliação da qualidade do MUCS obtido.

É bastante importante começar por definir o propósito do MUCS que se vai criar. Desta forma, é expectável que o MUCS cumpra determinados requisitos, como um valor mínimo de exatidão global e que siga uma determinada classificação. Devemos seleccionar a área que se pretende cobrir e a altura no tempo a que se referirá o MUCS. Depois de seleccionada a área de estudo, dever-se-á obter dados sobre a região que permitam conhecer a área e ajudar na criação do MUCS. Estes dados, usualmente imagens de satélite, caracterizam a região num ou mais instantes temporais devendo assim estar de acordo com o período de tempo definido.

Depois de obtidos os dados que caracterizem o terreno, devem-se identificar as possíveis *features*, isto é, as variáveis que caracterizam cada pixel. Usualmente estas *features* são as bandas das imagens de satélite, todas ou uma seleção.

Quando não se tem conhecimento prévio de áreas no terreno classificadas, o uso de classificadores não supervisionados permitem obter uma primeira classificação dos dados, necessitando apenas de alguns parâmetros de entrada, como o número de classes que se pretendem usar para dividir os dados. Um dos problemas que pode advir desta abordagem é a não correspondência direta das classes obtidas com as classes pretendidas. Assim, conhecendo previamente uma classificação parcial dos dados em função das classes de interesse, é possível usar classificadores supervisionados para fazer a classificação da área de interesse. Desta forma, é esperado que os resultados obtidos com um classificador supervisionado sejam melhores que os obtidos com um classificador não-supervisionado.

Definamos como áreas de treino as áreas ocupadas pelos pixels onde são conhecidas as classes de interesse. Quando a estes pixels se fazem corresponder variáveis que os caracterizam, *features*, obtêm-se os dados de treino.

Para se criar os dados de treino selecciona-se uma amostra de pixels, e é feita uma correspondência de um vetor multivariável a uma das classes (no caso de se usar um classificador rígido), ou a várias classes com proporções associadas (no caso do classificador a usar seja não rígido).

Depois de se criar os dados de treino, procede-se à classificação de imagens, treinando e aplicando o classificador a imagens de satélite, obtendo assim um MUCS. Depois de criado o mapa é necessário avaliar a qualidade do resultado obtido. Desta forma, valida-se o MUCS obtido comparando-o com dados de referência, podendo ser feita a comparação de forma global como para cada classe.

2.2. Missão Sentinel-2

A missão S2, desenvolvida pela Agência Espacial Europeia (também conhecida por ESA), é uma família de satélites cujos sensores instalados recolhem imagens nas bandas do visível e do infravermelho do espectro eletromagnético. Esta missão enquadra-se no programa *Copernicus* com o objetivo de monitorizar a Terra. Na constituição do seu segmento espacial tem-se um par de satélites idênticos posicionados numa mesma órbita e afastados entre si 180°, tendo uma cobertura terrestre entre as latitudes 56°S e 84°N. Em conjunto, os dois satélites revisitam um mesmo local ao longo do equador a cada 5 dias, demorando o dobro do tempo individualmente. Em latitudes intermédias o seu tempo de revisita pode demorar entre dois a três dias (Delwart, 2015).

Ambos os satélites têm na sua constituição um sensor multiespectral MSI (*MultiSpectralInstrument*) com treze bandas espectrais (Tabela 1): quatro bandas com resolução espacial de 10m, seis bandas com resolução espacial de 20m e três bandas com resolução espacial de 60m (Drusch et al., 2012; Zhang et al., 2019).

Bandas	Nome	Comprimento de onda central (nm)	Largura de banda (nm)	Resolução espacial (m)
b1	<i>Coastal aerosol</i>	443	20	60
b2	<i>Blue</i>	490	65	10
b3	<i>Green</i>	560	35	10
b4	<i>Red</i>	665	30	10
b5	<i>Rededge</i>	705	15	20
b6	<i>Rededge</i>	740	15	20
b7	<i>Rededge</i>	783	20	20
b8	<i>NIR</i>	842	115	10
b8A	<i>Narrow NIR</i>	865	20	20
b9	<i>Water vapour</i>	945	20	60
b10	<i>SWIR - Cirrus</i>	1375	30	60
b11	<i>SWIR</i>	1610	90	20
b12	<i>SWIR</i>	2190	180	20

Tabela 1: Bandas multiespectrais do Sentinel-2.

Esta missão dedica-se assim essencialmente a serviços de: monitorização do meio terrestre, por exemplo a partir da criação de MUCS; de suporte e emergências; relativos à segurança, por exemplo, de apoio à vigilância marítima e de fronteiras; e de monitorização de alterações

climáticas. Os produtos disponibilizados pela missão S2 estão categorizados em níveis, não estando todos disponíveis para os utilizadores (Delwart, 2015).

O produto de Nível 0, não disponível para os utilizadores, corresponde a dados de imagem comprimidos que são a base para a produção dos produtos de Nível 1. Os produtos de Nível 1, dividem-se em três subprodutos, 1A, 1B e 1C. O produto 1A, não disponível para utilizadores, é obtido pela descompressão do produto Nível 0, onde é aplicado um modelo geométrico que permite qualquer pixel da imagem ser localizado na superfície terrestre. Para o produto 1B, disponível para os utilizadores, consiste em imagens corrigidas radiometricamente para os valores de radiância no topo da atmosfera. O produto 1C resulta do uso de modelos digitais de elevação para obter as coordenadas cartográficas projetadas da imagem, apresentando os valores de refletância por cima da atmosfera. Quanto ao Nível 2, o subproduto 2A é um derivado do produto 1C, fornecendo assim refletâncias abaixo da atmosfera corrigidas em ortoimagens.

Os sensores dos satélites apresentam uma resolução radiométrica de 12 bits, significando que podem discriminar 4096 valores diferentes (variando de 0 a 4095). Posteriormente, estas medições são convertidas em refletâncias e armazenadas em imagens de inteiros de 16 bits.

2.3. Fontes de Dados Classificados

A quantidade de dados nas mais variadíssimas áreas tem crescido de forma exponencial e consequentemente a necessidade de saber como explorá-los também tem crescido. Também aos dados geográficos, isto é, dados que têm localizações geográficas associadas, se aplica este facto. Alguns exemplos deste tipo de dados são os MUCSs criados por entidades oficiais e a informação geográfica voluntária criada por qualquer cidadão.

Dentro do conjunto de MUCSs existentes, têm-se por exemplo a Carta de Ocupação e Uso do Solo (COS) e o Urban Atlas (UA). Ambos os produtos são oficiais. No caso da COS, esta é produzida para Portugal inteiro enquanto que o UA é existe apenas para cidades pan-europeias com uma população superior a 100.000 habitantes.

Para a informação geográfica voluntária (IGV), pode-se destacar os dados disponibilizados pelo OpenStreetMap (OSM). Estes dados são criados por voluntários e a sua classificação não engloba só classes de cobertura e uso do solo. O objetivo principal desta plataforma é mapear a Terra por inteiro.

Nas subsecções seguintes são apresentadas informações mais pormenorizadas sobre cada um dos dados apresentados anteriormente: COS, UA e OSM.

2.3.1. Carta de Ocupação e Uso do Solo

A COS corresponde a um produto de cartografia de ocupação e uso do solo, cuja produção é da responsabilidade da Direção Geral do Território (DGT). Esta cartografia constitui uma série temporal de cartas, com referências aos anos de 1995, 2007, 2010, 2015 e 2018, que caracterizam a ocupação e uso do solo do território nacional (Portugal Continental e Arquipélagos da Madeira e Açores). Com a atualização das cartas, para que haja consistência temporal e geométrica entre as várias cartas, as anteriores sofrem modificações, criando-se assim novas versões das mesmas, onde também são corrigidas áreas incorretamente classificadas (Direção-Geral do Território, 2018, 2019).

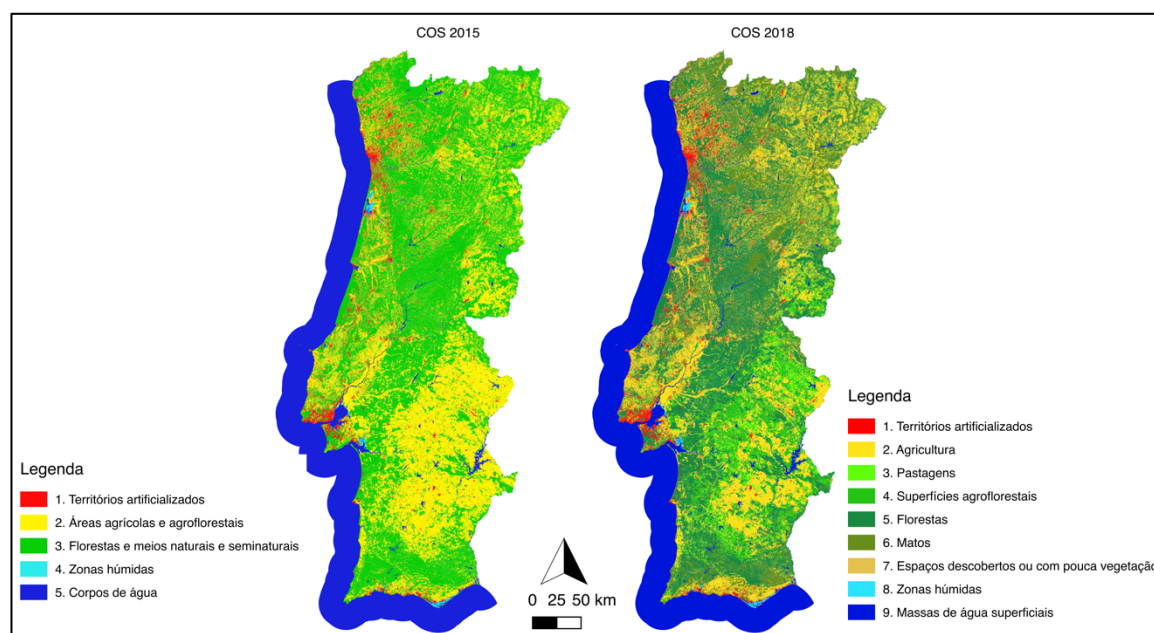


Figura 1: Nível 1 das COS 2015 e COS 2018 para Portugal Continental

Todas as versões produzidas da COS são cartografias de polígonos, onde cada polígono representa a delimitação de uma área do terreno pertencente a uma determinada classe de ocupação ou uso do solo de forma homogénea. A área de cada polígono tem de ser superior à unidade mínima cartográfica (UMC) definida, isto é, a área mínima que um polígono pode ter, que é 1ha para todas as versões da COS, com distância mínima entre linhas de 20m. A classe de ocupação e uso do solo atribuída a cada polígono tem de ocupar pelo menos 75% da sua área. No caso de áreas com dimensão inferior à UMC, estas são generalizadas de acordo com um conjunto de regras definidas.

A produção das várias cartas foi feita com base em fotointerpretação de imagens aéreas ortoretificadas com grande resolução espacial e com referência aos respetivos anos. No caso da COS mais recente, 2018, foi produzida por atualização da anterior, de 2015.

A nomenclatura das várias versões segue uma estrutura hierárquica por níveis, sendo o primeiro nível o mais generalizado.

Quanto ao nível mais geral da nomenclatura, até 2015, era composto por 5 classes (Figura 1), sendo elas (Direção-Geral do Território, 2018):

1. *Territórios artificializados*: superfícies artificializadas ou ajardinadas, destinadas a atividades que envolvam a intervenção humana, como edifícios, áreas industriais e estradas;
2. *Áreas agrícolas e agroflorestais*: áreas utilizadas para agricultura, incluindo sistemas agroflorestais;
3. *Florestas e meios naturais e seminaturais*: áreas com florestas, vegetação arbustiva e herbácea, incluindo áreas naturais com pouca ou nenhuma vegetação;
4. *Zonas húmidas*: zonas húmidas interiores ou litorais, cobertas temporariamente ou permanentemente por água;
5. *Corpos de água*: superfícies de água doce ou salgada.

A COS 2018, passou a seguir uma nova nomenclatura com base na anterior, sendo o primeiro nível dividido em 9 classes de ocupação e uso do solo (Figura 1), sendo elas (Direção-Geral do Território, 2019):

1. *Territórios artificializados*: áreas destinadas a atividades que envolvam a intervenção humana, como edifícios, áreas industriais e estradas;
2. *Agricultura*: áreas utilizadas para agricultura;
3. *Pastagens*: áreas ocupadas com vegetação essencialmente do tipo herbácea ocupando uma área superior a 25% da superfície;
4. *Superfícies agroflorestais*: associação vertical numa mesma parcela de culturas temporárias e/ou pastagens e/ou culturas permanentes com espécies florestais com um grau de coberto superior ou igual a 10%;
5. *Florestas*: áreas com uso florestal, ocupadas por árvores florestais ou temporariamente desarborizados. As árvores devem ter uma altura superior ou igual a 5 metros e em conjunto um grau de coberto superior ou igual a 10%;
6. *Matos*: áreas naturais com vegetação espontânea, com coberto arbustivo superior ou igual a 25%;
7. *Espaços descobertos ou com pouca vegetação*: áreas naturais com pouca ou nenhuma vegetação. Estas áreas devem incluir rocha nua, praias e areais e vegetação esparsa, ocupando no máximo 25% da área;
8. *Zonas húmidas*: áreas interiores ou litorais, cobertas temporariamente ou permanentemente por água;
9. *Massas de água superficiais*: superfícies de água doce ou salgada.

Estes produtos seguem assim um modelo de dados vetorial, disponibilizados no formato *ESRI shapefile* no sistema de georreferência ETRS/PT-TM06. É feita a associação de tabelas alfanuméricas com atributos que caracterizam cada um dos polígonos que delimitam as áreas. Dos atributos destacam-se o identificador, que permite identificar de forma única cada um dos polígonos, e um código que caracteriza a classe de ocupação e uso do solo referente ao nível mais detalhado a que o respetivo polígono pertence. Quanto à qualidade destas cartas, até 2015, apresentam exatidões temáticas superiores ou iguais a 85% e exatidões posicionais melhores ou iguais a 5,5m (Direção-Geral do Território, 2018, 2019). Para a COS 2018, até dezembro de 2019, esta encontrava-se em processo de avaliação da exatidão temática.

2.3.2. Urban Atlas

O UA é um produto criado para mapear áreas urbanas pan-europeias, enquadrado no programa *EU Copernicus* e fornece dados de uso e cobertura do solo. Os dados disponibilizados são apoiados por observações da Terra e por outros dados de referência, como Comercial Off-The-Self (COTS), OpenStreetMap (OSM) e mapas topográficos (Copernicus, 2016).

Atualmente, o UA apresenta três produtos com referência aos anos 2006, 2012 e 2018. À data, o produto mais recente encontra-se disponível, mas não validado. Todos estes produtos são disponibilizados em formato vetorial e, tal como na COS, cada polígono representa a delimitação de uma área com classificação homogénea quanto à classe de uso e cobertura a que pertence, apresentando áreas superiores UMC definida.

Em Portugal (Figura 2), este produto relativo ao ano de 2012 é criado apenas para as cidades de: Aveiro, Braga, Coimbra, Faro, Funchal, Guimarães, Lisboa, Porto, Póvoa de Varzim, Viana do Castelo e Viseu. Para 2018, este produto encontra-se disponível para todas as cidades anteriores exceto para Braga, encontrando-se ainda em produção em agosto de 2020 (<https://land.copernicus.eu/local/urban-atlas/urban-atlas-2018?tab=mapview>).

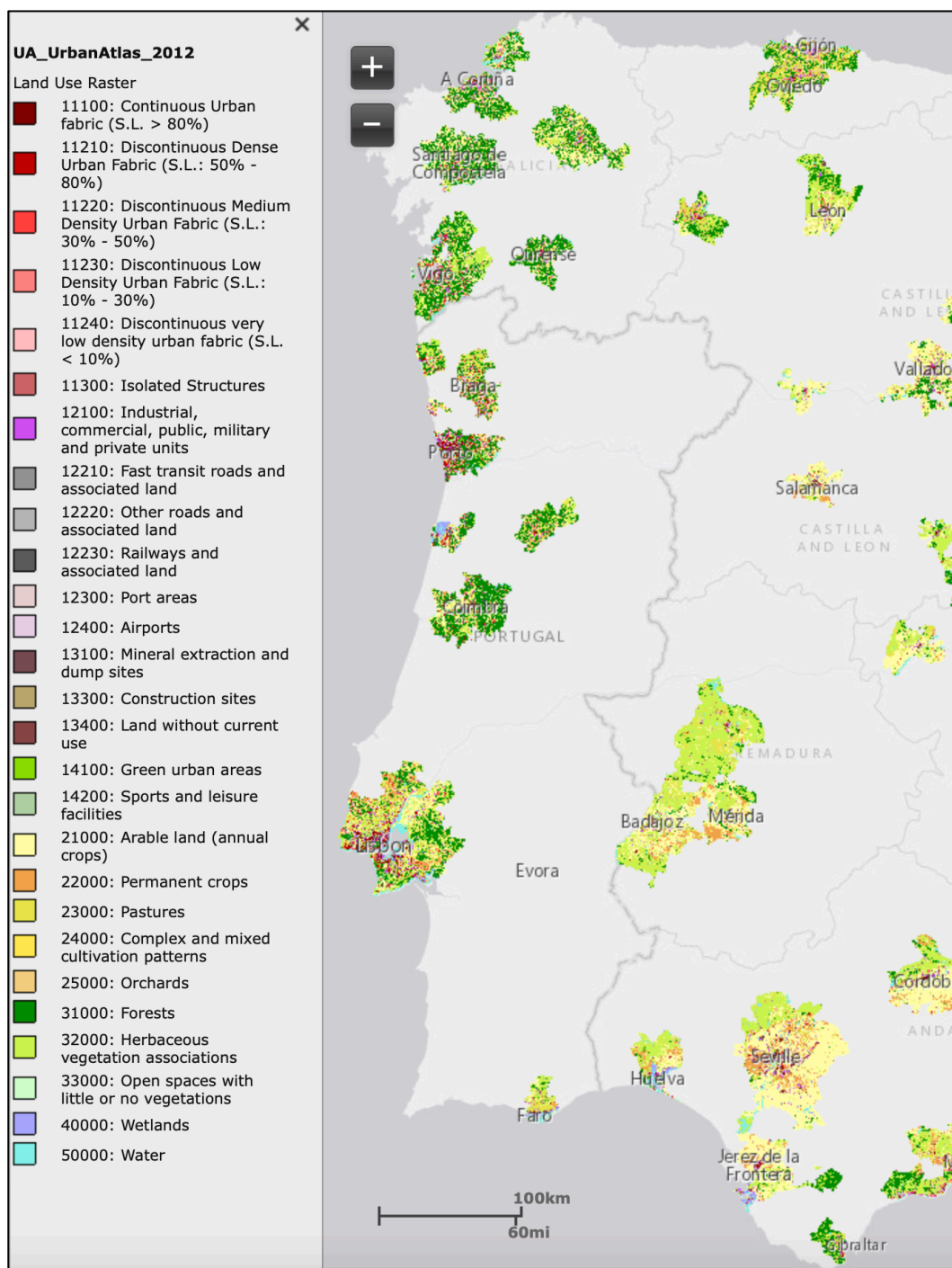


Figura 2: Cidades abrangidas pelo UA 2012 para Portugal Continental e para parte de Espanha (Fonte: <https://land.copernicus.eu/local/urban-atlas/urban-atlas-2012>).

De forma semelhante à COS, a nomenclatura utilizada no UA também segue uma estrutura hierárquica de 4 níveis, sendo assim atribuído a cada polígono um código de 5 dígitos de forma a identificar a classe a que o polígono pertence. No nível mais pormenorizado, o UA de 2012 e de 2018 contam com 27 classes de ocupação e uso do solo e adicionalmente 2 classes relativas a zonas não classificadas. Para o primeiro nível, nível mais generalizado, os UA de 2012 e de 2018 apresenta 5 classes com a seguinte nomenclatura:

1. *Territórios artificializados*: superfícies que sofreram a intervenção humana sem solo agrícola. Esta classe inclui todo o tipo de estruturas artificiais e todas as superfícies associadas a estes, com ou sem vegetação;
2. *Áreas agrícolas*: áreas utilizadas para agricultura, onde se inclui solos aráveis, culturas permanentes, pastagens, pomares e áreas agroflorestais;
3. *Áreas naturais e seminaturais*: áreas florestais, áreas com vegetação do tipo herbáceo e espaços com pouca ou nenhuma vegetação;
4. *Zonas húmidas*: áreas com interiores e litorais cobertas temporariamente ou permanentemente com água doce ou salgada;
5. *Água*: corpos de água e cursos de água.

Para o UA de 2012 e de 2018, no caso da primeira classe, esta apresenta uma UMC de 0,25ha e exatidão temática não inferior a 85%. Nas restantes classes, a UMC é de 1ha e apresenta uma exatidão temática não inferior a 80%. Assim, de modo geral, este produto apresenta uma exatidão temática maior ou igual a 80% (Gallaun & Hirschmugl, 2017).

2.3.3. Informação Geográfica Voluntária - OpenStreetMap

A IGV é um tipo de conteúdos georreferenciados criados pelos cidadãos, em geral de forma voluntária. Com a evolução das tecnologias Web, a interação do utilizador com estas tecnologias tem vindo a melhorar, havendo uma interação mais intuitiva e fácil para o utilizador disponibilizar este tipo de informação.

A IGV pode-se dividir em dois tipos: a implícita e a explícita (Senaratne et al., 2017). No caso da implícita, o utilizador associa uma localização georreferenciada a conteúdos multimédia, como imagens, textos e vídeos. Como exemplo pode-se indicar a publicação de uma fotografia no Instagram a que é atribuída uma localização, ou quando se cria uma publicação no Twitter e se associa uma localização geográfica. Para a IGV explícita, o utilizador está focado em atividades de mapeamento, contribuindo com dados que têm a si associados conteúdos geográficos, como por exemplo, a criação de geometrias (do tipo ponto, linha e polígono) no OSM e no Google.

O papel da IGV tem sido importante, particularmente em situações de desastres. A recolha de informação geográfica em tempo quase real tem permitido uma resposta mais eficiente na gestão de recursos para o solucionamento destes problemas. Associado a este tipo de informação, é importante a avaliação da sua qualidade uma vez que pode ser criada por qualquer pessoa. Como já demonstrado em trabalhos de Chunara et al. (2012), Fuchs et al. (2013), MacEachren et al. (2011)

e Sakaki et al. (2010), quanto ao uso de IGV e potencialidades deste tipo de informação tem vindo a aumentar ao longo do tempo (Senaratne et al., 2017). Em Goodchild & Li (2012) são apresentadas três possíveis formas de avaliar a qualidade da IGV: aproximações *crowd-sourcing*, social e geográfica.

Na aproximação *crowd-sourcing*, a ideia principal é qualquer pessoa, sem ter em conta a sua formação, possa validar e corrigir erros na IGV criada. Com esta aproximação, erros geográficos, associados ao posicionamento geográfico, poderão prevalecer, uma vez que, por exemplo, os utilizadores podem ou não ter formação na área e as imagens aéreas utilizadas possam não estar ortorretificadas.

Para a aproximação social, aos utilizadores voluntários são atribuídas reputações consoante o trabalho prestado. Assim, existe uma hierarquização nos utilizadores, esperando que utilizadores mais assíduos e com trabalho avaliado positivamente tenham conteúdos de maior confiança.

Na aproximação geográfica, com o conhecimento geográfico e de legislações, são aplicadas regras de verificação de forma a validar a IGV criada e editada. Por exemplo, numa zona classificada como água não podem existir edifícios.

Um dos projetos colaborativos que permite a inserção/criação/edição de IGV é o OSM. Este projeto foi criado no *University College London* em 2004 por Steve Coast, contando atualmente com alguns milhões de colaboradores por todo o mundo, normalmente mais ativos em áreas urbanas. O OSM contém dados vetoriais geoespaciais, ou seja, geometrias do tipo ponto, linha e polígono. Além destas geometrias tem-se relações lógicas entre geometrias e relações entre relações. A estas geometrias e relações são associados atributos ou *tags* que caracterizam os elementos do OSM (Calazans Campelo et al., 2017).

Neste projeto, além de um utilizador poder contribuir com IGV, também tem acesso à mesma, sendo os dados de livre acesso. Desta forma, o OSM tem como objetivos o mapeamento do mundo inteiro e a disponibilização gratuita dessa informação, contrariando assim a capitalização da informação geográfica disponibilizada por outras plataformas que não permitem o acesso livre à informação, acabando por esta não estar disponível para muitos utilizadores individuais, pequenas empresas e a organizações comunitárias.

Com o aumento da popularidade do OSM, o número de utilizadores tem aumentado e permitido que a quantidade de IGV tenha aumentado, em conjunto com a sua qualidade. A partir dos dados existentes nesta plataforma podem-se criar MUCSs a partir da conversão direta dos dados do OSM em classes de uso e ocupação do solo. Desta forma, existe um grande interesse na comparação destes dados com fontes oficiais de dados (Borghuis et al., 2007; Haklay, 2010). As *tags* associadas às geometrias no OSM são constituída por uma chave e um valor, por exemplo, *building=residential* (lista completa em https://wiki.openstreetmap.org/wiki/Map_Features), sendo a partir destas que se pode extrair informação de ocupação e uso do solo (Calazans Campelo et al., 2017).

2.4. Classificadores de Dados

O processo de classificação de dados corresponde a atribuir aos dados etiquetas que categorizam os mesmos. Desta forma, obtêm-se dados com a mesma etiqueta que em conjunto constituem uma classe. Pode-se assim definir uma classe como um conjunto de dados que têm características comuns.

Um exemplo prático, do dia-a-dia, de uma classificação, é como distinguimos um conjunto de peças de fruta, por exemplo tendo um cesto de fruta com bananas, pêssegos e alperces. Ao distinguirmos as peças de fruta em “bananas”, “pêssegos” e “alperces” estamos a criar um processo de classificação. Mas para podermos criar este processo de classificação precisamos de recorrer às características de cada uma das frutas, designadas por *features*. Neste caso, pode-se pensar nalgumas características que pudessem permitir esta distinção, como a cor exterior, a forma e a textura. Estas são algumas das características que se poderiam considerar. Algumas delas, por si só permitiam distinguir facilmente cada uma das peças de fruta, por exemplo, usando só a forma da peça da fruta, iríamos conseguir facilmente separar as bananas do resto das peças de fruta, mas os alperces e os pêssegos têm formas muito semelhantes, havendo uma certa confusão no processo de classificação se não considerarmos mais características. Se usássemos por exemplo a cor exterior e a forma, este processo de classificação iria conseguir, à partida, separar sem qualquer problema cada uma das peças de frutas. Usando mais *features*, estas poderiam ser redundantes no processo de classificação ou então conseguir garantir que o processo de classificação iria ter melhor performance ou até mesmo baralhar o processo de classificação.

Define-se um padrão como sendo um conjunto de variáveis que caracterizam um indivíduo. O reconhecimento de padrões é a ciência que estuda a descrição e a classificação dos objetos, ou seja, o estudo de padrões. Assim, num processo de reconhecimento de padrões pretende-se separar e/ou classificar os objetos em categorias. Consoante o classificador utilizado, os dados de entrada e de saída poderão ser diferentes.

Quando a radiação emitida pelo sol incide sobre materiais, interage com estes, podendo interagir das seguintes formas (Ihlen, 2019): transmissão (a energia passa sobre o material); absorção (a energia é perdida sobre o objeto através de reações moleculares ou eletrónicas); reflexão (a energia é reenviada para o espaço com iguais ângulos de reflexão e incidência); dispersão (a energia é reenviada para o espaço com uma direção aleatória); e/ou emissão (a energia é emitida pelo corpo, havendo primeiro uma absorção da energia solar e conseqüentemente um aquecimento do corpo).

Cada corpo interage de maneira diferente com a radiação solar, apresentando assim uma assinatura espectral, isto é, para cada comprimento de onda apresenta uma percentagem de refletância característica. Na Figura 3 tem-se o exemplo de várias curvas espectrais para corpos

diferentes. Desta forma, é expectável que para dois corpos semelhantes, para um mesmo comprimento de onda, a refletância espectral seja a mesma.

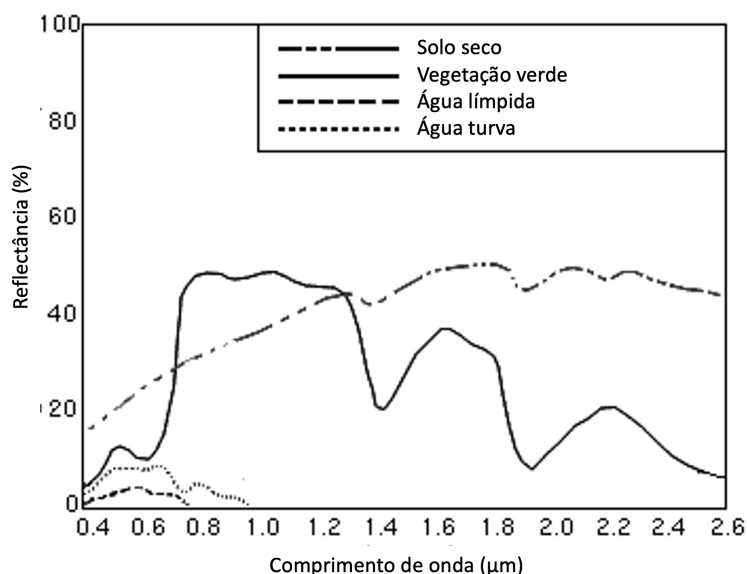


Figura 3: Curvas de refletância da água límpida, água turva, da vegetação e de solo seco (adaptado de Dalezios, 2014).

Em cada visita pelo satélite, o sensor multiespectral mede valores de radiância a partir da superfície terrestre para cada uma das bandas espectrais. A partir da radiância medida e da irradiância é determinada a refletância, a proporção entre a energia que é refletida e/ou emitida. Desta forma, um pixel pode ser caracterizado pelas respostas espectrais nos vários comprimentos de onda em que o sensor do satélite opera.

Os classificadores podem ser classificados quanto a usarem dados de treino ou não. Quando um classificador é desenvolvido a partir de dados classificados, dados de treino, denominam-se estes como classificadores supervisionados. Neste tipo de classificadores os dados de treino permitem que o classificador reconheça novos padrões a partir das características dos padrões classificados previamente. Um classificador não supervisionado não necessita de dados de treino como entrada, mas necessita de parâmetros pré-definidos, como, por exemplo, o número de classes. A partir dos dados a classificar, o classificador organiza os dados em grupos, (*clusters*), que são divididos por terem combinações de características semelhantes, podendo ou não ter qualquer significado físico esta divisão dos padrões (Sharma & Kaur, 2013).

Outra classificação dos classificadores é se estes são rígidos ou não rígidos (Hu et al., 2013). No caso dos classificadores rígidos, a cada unidade que se pretende classificada é atribuída uma e uma só classe, sendo assim criado um só mapa. Para o segundo tipo de classificadores, os não rígidos, a cada unidade é atribuída uma probabilidade, possibilidade ou um grau de pertença à classe (Bede, 2013; Kuncheva, 2000), sendo assim criado um mapa por classe.

Antes de qualquer trabalho a ser realizado, é importante ter em conta que não existe um classificador perfeito, capaz de realizar qualquer classificação com uma qualidade de 100%. Neste trabalho, irá ser usada a versão rígida do classificador supervisionado “*Random Forest*”.

O classificador *Random Forest* baseia-se no conceito de florestas aleatórias, introduzido inicialmente por Leo Breiman (Breiman, 2001). A nível computacional, este classificador é bastante atrativo porque se baseia tanto em regressões como na classificação multiclasse, é relativamente rápido de ser desenvolvido e aplicado, e é facilmente aplicado a problemas de grande dimensão, ou seja, a padrões com uma grande quantidade de características (*features*) (Cutler et al., 2012).

No classificador *Random Forest* é criado um conjunto de árvores de decisão aleatórias a partir de um conjunto de dados de treino e aplicadas sobre os dados a classificar. Numa árvore de decisão criam-se critérios a partir de uma ou mais *features* e tomam-se decisões até se obter uma resposta. A decisão da classificação utilizando o classificador *Random Forest* é assim determinada por um critério de maioria, considerando a resposta com maior frequência.

Utilizando o exemplo anterior da cesta de frutas, pegando num pêsego, pode-se começar com a pergunta “A forma é de uma banana?”, a resposta seria “Não” e ainda não tínhamos a resposta de que peça de fruta seria. O processo continuaria e poder-se-ia colocar a pergunta “Tem cor exterior avermelhada?”, e a resposta também seria “Sim” e teríamos assim a resposta que a peça de fruta em questão era um pêsego. Este exemplo pode ser representado por uma árvore de decisão apresentada na Figura 4. Utilizando várias árvores de decisão poderíamos obter várias classificações, entrando assim o critério de maioria para decisão final de classificação.

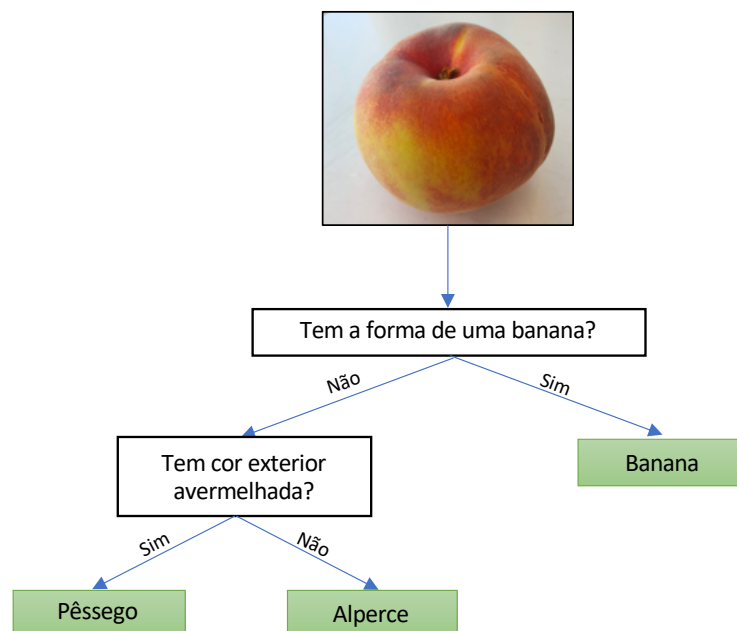


Figura 4: Exemplo de árvore de decisão. (Foto cedida por Tiago Rocha)

Este classificador tem sido usado em bastantes trabalhos de investigação, como em Htitiou et al. (2019), apresentando bons resultados em classificação de imagens de satélite. Em *Python*, este classificador encontra-se implementado na biblioteca *scikit-learn*.

A utilização deste classificador é feita em três etapas: definição do classificador, treino do classificador e aplicação do classificador. Quanto à definição do classificador, este pode ser utilizado de imediato com as definições padrão, nomeadamente o número de árvores de decisão (*n_estimators=100*) e a profundidade máxima de cada árvore (*max_depth=None*). Para treinar o classificador apenas é necessário como parâmetros de entrada um *array* com os dados de treino e um *array* com a respetiva classificação. Desta forma constrói-se uma floresta de árvores de decisão que são usadas na classificação dos dados. Para a aplicação do classificador utiliza-se como parâmetro de entrada uma matriz com os dados a classificar, tendo como retorno um *array* com a classificação obtida.

Uma boa prática nos processos de classificação é a análise das variáveis que caracterizam os padrões, por exemplo, avaliando a separabilidade de cada classe em relação às restantes. Para a avaliação da separabilidade das classes, para cada classe usa-se uma métrica que avalie a distância entre o *cluster* constituído pela classe em questão e o *cluster* constituídos pelos dados pertencentes às restantes classes. Existem muitas medidas para avaliar a separabilidade, nomeadamente, a divergência, a divergência transformada e a distância de Jeffries-Matusita (Kavzoglu & Mather, 2020). Neste trabalho usar-se-á a distância de Bhattacharya que será explicada posteriormente na secção 3.4.

2.5. Índices Radiométricos

Pode-se definir um índice radiométrico como uma banda adicional, obtida por operações matemáticas a partir de um conjunto de bandas das imagens multiespectrais. Os índices radiométricos têm uma grande importância na monitorização do solo e cada um tem a sua aplicação. Consoante o índice a que nos referimos, podemos aplicá-los a cobertura do solo diferentes. Por exemplo, temos índices para monitorizar a água (por exemplo, o NDWI) e temos índices para monitorizar a vegetação (por exemplo, o NDVI).

Neste trabalho foram utilizados quatro índices radiométricos diferentes: NDVI (*Normalized Difference Vegetation Index*), NDWI (*Normalized Difference Water Index*), NDBI (*Normalized difference Built-up Index*) e SAVI (*Soil Adjusted Vegetation Index*). Estes índices utilizados, são apenas alguns dos muitos que existem e são explicados nas subsecções seguintes. Estas novas

bandas, poderão ser assim usadas como bandas adicionais na classificação de imagens e/ou como informações adicionais para a filtragem dos dados de treino. A decisão no uso destes índices está no facto de estarem associados às principais classes de cobertura do solo: vegetação, água e edificado.

2.5.1. NDVI

O Índice de Diferença Normalizada de Vegetação (NDVI – Normalized Difference Vegetation Index) quantifica a diferença entre as respostas espectrais com radiação infravermelha próxima e a luz vermelha. Estas duas radiações têm comportamentos opostos quando incidem sobre vegetação. Enquanto a luz vermelha é absorvida pela vegetação, a radiação infravermelha é refletida. Este índice é assim calculado a partir da fórmula (1), onde *NIR* corresponde à banda do infravermelho próximo e *Red* corresponde à banda do infravermelho próximo.

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

Este índice varia entre -1 e 1. Quando a vegetação é saudável, os vários tipos de clorofila refletem mais radiação no infravermelho próximo e no verde, absorvendo mais luz vermelha e azul. Desta forma, como na zona visível do espectro é apenas refletida a luz verde, os nossos olhos veem verde. A partir da fórmula de cálculo do NDVI, se houver baixa refletância na zona do vermelho e alta refletância na zona NIR (Luo et al., 2019), o NDVI tomará valores mais altos, significando que se está na presença de vegetação saudável e densa. Assim, este índice é utilizado, por exemplo, na monitorização e deteção de vegetação (Albarakat & Lakshmi, 2019).

Utilizando imagens de S2, este índice é assim calculado a partir das bandas 4 e 8, tendo assim a fórmula (2), onde *b8* corresponde à banda 8 do S2 e *b4* corresponde à banda 4 do S2.

$$NDVI = \frac{b8 - b4}{b8 + b4} \quad (2)$$

2.5.2. NDWI

De forma semelhante ao NDVI, o Índice de Diferença normalizada de água (NDWI – Normalized Difference Water Index) quantifica a diferença entre as respostas espectrais à radiação verde e à radiação infravermelha próxima, sendo assim calculado pela fórmula (3), onde *Green* corresponde à banda da luz verde e *NIR* corresponde à banda do infravermelho próximo.

$$NDWI = \frac{Green - NIR}{Green + NIR} \quad (3)$$

Este índice, tal como o NDVI, varia entre -1 e 1. Como a água absorve a maior parte da radiação infravermelha, ou seja, há uma baixa refletância da radiação do infravermelho próximo, quando este índice toma valores positivos estamos na presença de corpos de água (McFeeters, 2013) e em caso de negativos outro tipo de cobertura de superfície.

Utilizando imagens de S2, este índice é assim calculado a partir das bandas 3 e 8, tendo assim a fórmula (4), onde *b3* corresponde à banda 3 do S2 e *b8* corresponde à banda 8 do S2.

$$NDWI = \frac{b3 - b8}{b3 + b8} \quad (4)$$

2.5.3. NDBI

O NDBI (Normalized Difference Built-up Index), Índice de Diferença Normalizada de Construções, quantifica a diferença entre as respostas espectrais com a radiação infravermelha adjacente à radiação infravermelha próxima (SWIR) e a radiação infravermelha próxima (NIR), sendo assim calculado pela fórmula (5), onde *SWIR* corresponde à banda da radiação infravermelha adjacente à radiação infravermelha próxima e *NIR* corresponde à banda do infravermelho próximo.

$$NDBI = \frac{SWIR - NIR}{SWIR + NIR} \quad (5)$$

Este índice, tal como os anteriores, também varia entre -1 e 1, sendo utilizado para a deteção de construções artificiais (He et al., 2010). No caso de se estar na presença de construções, é esperado que o valor do índice NDBI seja positivo (Karanam, 2018), e negativo caso contrário. De forma análoga à análise dos índices anteriores, significa assim que as construções absorvem uma maior quantidade de radiação NIR comparativamente à radiação SWIR, ou seja, a refletância da radiação NIR será inferior à refletância da radiação NIR.

Utilizando imagens de S2, este índice é assim calculado a partir das bandas 3 e 8, tendo assim a fórmula (6), onde $b1$ corresponde à banda 11 do S2 e $b8$ corresponde à banda 8 do S2.

$$NDBI = \frac{b11 - b8}{b11 + b8} \quad (6)$$

2.5.4. SAVI

O Índice de Vegetação Ajustado ao Solo (SAVI – Soil Adjusted Vegetation Index) quantifica uma relação entre as respostas espectrais da radiação infravermelha próxima (NIR) e a radiação vermelha utilizando um fator pré-definido, sendo calculado pela fórmula (7), onde NIR corresponde à banda da radiação do infravermelho próximo, Red corresponde à banda da radiação vermelha e L corresponde ao fator de correção, sendo tipicamente 0.5.

$$SAVI = \frac{NIR - Red}{NIR + Red + L} (1 + L) \quad (7)$$

Este índice, calculado de forma semelhante ao NDVI, usa um fator de correção para determinar a influência da refletância do solo quando a vegetação é pouco densa (Huete, 1988).

Utilizando imagens de S2, este índice é assim calculado a partir das bandas 8 e 4, tendo assim a fórmula (8), onde $b8$ corresponde à banda 8 do S2, $b4$ corresponde à banda 4 do S2 e L corresponde ao fator de correção a definir.

$$SAVI = \frac{b8 - b4}{b8 + b4 + L} (1 + L) \quad (8)$$

2.6. Validação de MUCSs

A validação dos MUCSs consiste na avaliação da qualidade do processo de classificação. As formas mais usuais de validação da classificação de imagens de satélite, consistem na recolha de uma amostra de dados e na comparação com a realidade averiguada por fotointerpretação ou visitas de campo (Stehman, 2009; Stehman & Foody, 2019). Outra forma é avaliando a totalidade do mapa,

comparando a totalidade do resultado obtido pela classificação com um mapa considerado como referência, supostamente com maior qualidade do que o mapa a validar, usando assim todos os pixels classificados em vez de uma amostra. Usualmente, para se fazer a análise da qualidade de uma classificação, recorre-se a uma matriz de confusão. A partir desta, calculam-se índices de exatidão: nomeadamente a exatidão do utilizador e do produtor por classe, e a exatidão global. Na prática, a avaliação desta qualidade deverá refletir a proporção areal de cada classe na zona de estudo e deve ainda ser estatisticamente rigoroso (Stehman & Foody, 2019). Usando um mapa de referência totalmente classificado a proporção areal de cada classe já é verificada.

Na Tabela 2 é apresentada uma matriz de confusão para uma classificação com n classes. As linhas $i = 1, 2, \dots, n$ representam a classificação do mapa e as colunas $j = 1, 2, \dots, n$ a classificação de referência. Cada componente p_{ij} , $i, j = 1, 2, \dots, n$, corresponde à proporção de pixels classificados como pertencentes à classe i mas que na realidade pertencem à classe j . Quando $i = j$, tem-se a proporção de elementos classificados corretamente na classe em questão. Para $i \neq j$, a proporção obtida indica a proporção de pixels que são confundidos. As componentes p_{i*} e p_{*j} são as somas das proporções na linha i e coluna j , respetivamente, representando assim a proporção de pixels classificados no MUCS obtidos como pertencentes à classe i e a proporção de pixels que estão classificados como pertencentes à classe j no MUCS de referência, respetivamente.

Na linha marginal da matriz de confusão obtém-se a exatidão do utilizador, esta é obtida a partir do quociente entre a proporção de elementos classificados corretamente e a soma total das proporções na mesma linha. O complemento deste índice, isto é, $1 - p_{ii}/p_{i*}$ é denominado por erro de comissão para a classe i . Este tipo de erro ocorre quando o pixel é classificado como pertencente à classe em questão no mapa, mas na base de dados de referência pertence a outra classe diferente (Pontius Jr & Millones, 2011).

Na coluna marginal da matriz de confusão, tem-se a exatidão do produtor calculada a partir do quociente entre a proporção de elementos classificados corretamente e a soma total de proporções na respetiva coluna. O complemento deste índice, isto é, $1 - p_{jj}/p_{*j}$ é denominado por erro de omissão para a classe j , ocorrendo quando o pixel é classificado como pertence a uma classe que não a classe j , mas na base de dados de referência pertence à classe j (Pontius Jr & Millones, 2011).

Mapa Referência	C_1	C_2	...	C_j	...	C_n	Total	Exatidão do Produtor
C_1	p_{11}	p_{12}	...	p_{1j}	...	p_{1n}	p_{1*}	$\frac{p_{11}}{p_{1*}}$
C_2	p_{21}	p_{22}	...	p_{2j}	...	p_{2n}	p_{2*}	$\frac{p_{22}}{p_{2*}}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
C_i	p_{i1}	p_{i2}	...	p_{ij}	...	p_{in}	p_{i*}	$\frac{p_{ii}}{p_{i*}}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
C_n	p_{n1}	p_{n2}	...	p_{nj}	...	p_{nn}	p_{n*}	$\frac{p_{nn}}{p_{n*}}$
Total	p_{*1}	p_{*2}	...	p_{*j}	...	p_{*n}	1	
Exatidão do Utilizador	$\frac{p_{11}}{p_{*1}}$	$\frac{p_{22}}{p_{*2}}$...	$\frac{p_{jj}}{p_{*j}}$...	$\frac{p_{11}}{p_{*1}}$		

Tabela 2: Matriz de confusão para um problema de n classes.

A exatidão temática tem uma grande importância na decisão acerca do uso ou não de determinado mapa, ou seja, a sua adequabilidade para o objetivo em questão. Consoante a finalidade poderá ser necessário uma exatidão maior para certas classes, não acontecendo o mesmo para outras. De forma a avaliar globalmente a exatidão temática de um mapa, pode-se utilizar o índice de exatidão global obtido como a soma das componentes da diagonal principal da matriz, isto é:

$$E. \text{ Global} = \sum_{i=1}^n p_{ii} ,$$

com:

n = número de classes,

p_{ii} = proporção de pixéis classificados corretamente.

Assim, este índice não está associado a uma classe em particular e apenas fornece uma avaliação geral da qualidade do mapa, sendo recomendado complementar esta avaliação com as exatidões do utilizador e do produtor para cada uma das classes.

3. Metodologias de Criação Automática de Dados de Treino

Para a classificação automática supervisionada de imagens de satélite, a criação de dados de treino consiste na prática na delimitação de áreas no mapa (áreas de treino) correspondentes a cada uma das classes de interesse, às quais são depois associadas *features* que caracterizam estas áreas. Por exemplo, delimitando uma área que ocupe dois pixels, esta área irá dar origem a dois dados de treino, cada um referente a um pixel.

A delimitação de áreas de treino pode ser um processo demoroso e subjetivo. Assim, a criação automática de dados de treino é um processo desejado para quem tenha de fazer classificações automáticas supervisionadas de imagens de satélite.

Nas secções seguintes são apresentadas metodologias de criação de dados de treino a partir de fontes de dados classificados que posteriormente, capítulo 4, serão aplicadas a duas áreas de estudo, usando o mesmo processo de classificação e validação.

Para a criação dos dados de treino e seu uso na classificação automática de imagens do satélite S2, recorre-se à linguagem de programação *Python* e às seguintes bibliotecas já implementadas nesta linguagem: *numpy*, *osgeo.gdal*, *osgeo.ogr*, *osr*, *math*, *geopandas* e *fiona*.

Na Figura 5 apresentam-se as principais etapas da metodologia a usar. Em primeiro lugar, na secção 3.1, começa-se por definir a nomenclatura a usar e fazer a correspondência entre esta e as classes dos produtos que são usados (COS 2015, COS 2018, UA 2012 e OSM). De seguida, na secção 3.2, apresenta-se o processo de criação de grelhas constituídas por polígonos quadrangulares que delimitam os pixels das bandas de S2 com maior resolução espacial (10m). A cada polígono são associadas as classes existentes na área que ele ocupa, dos vários produtos usados. Com recurso às grelhas criadas, na secção 3.3, criam-se as áreas de treino que, em conjunto com as *features* selecionadas, constituem os dados de treino. De seguida, são apresentadas metodologias de filtragens dos dados de treino com recurso aos índices radiométricos, gerando-se assim novos dados de treino. Depois de criados os vários conjuntos de dados de treino, na secção 3.4 estes são analisados, sendo feita a avaliação da sua exatidão através da comparação com o mapa de referência e avaliando a separabilidade dos dados entre cada classe e as restantes. Se seguida, na secção 3.5, os dados de treino criados são usados para a classificação de imagens de satélite. A qualidade dos MUCSs resultantes é posteriormente avaliada, secção 3.6, através da comparação dos MUCSs obtidos com um mapa de referência com as características da COS 2018.

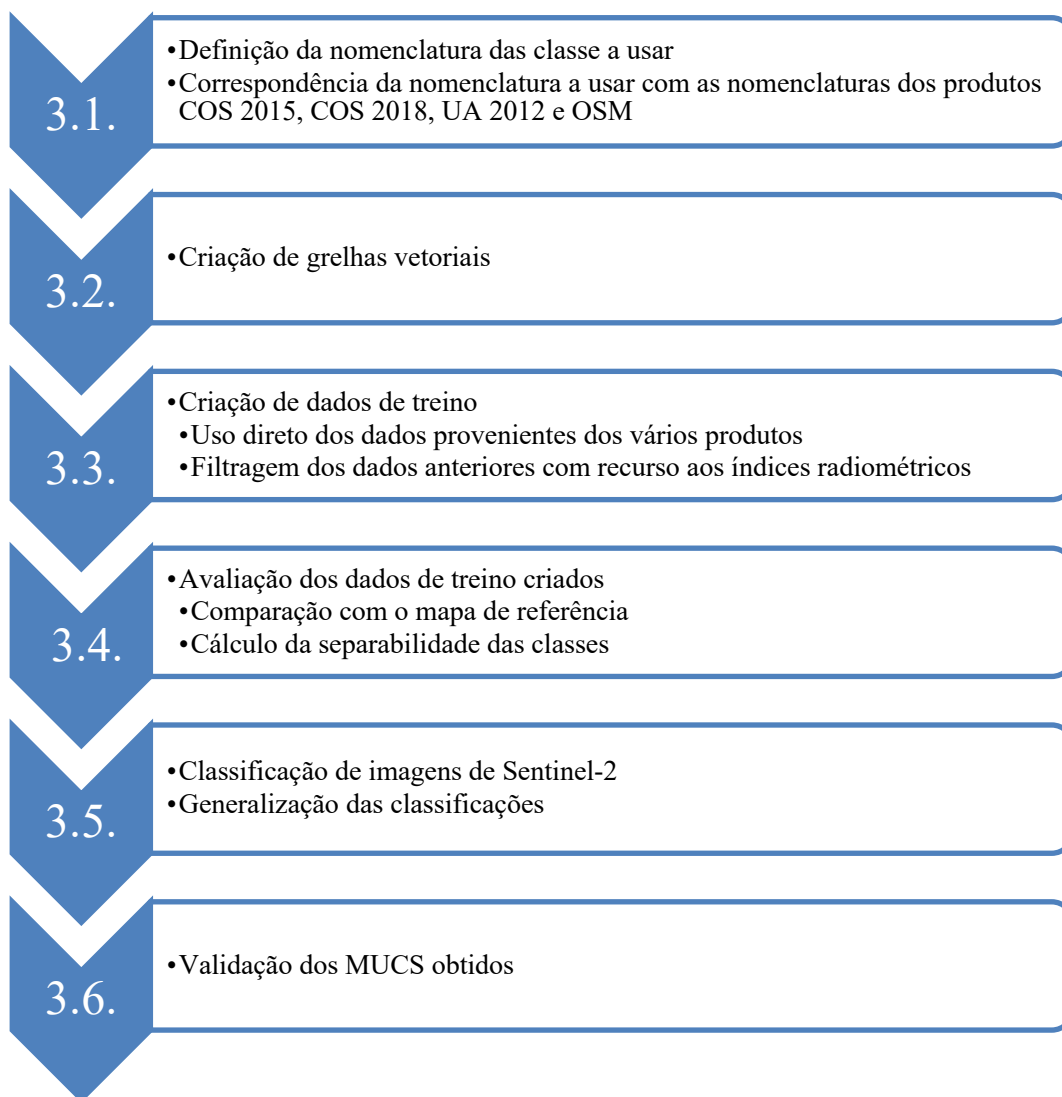


Figura 5: Etapas principais da metodologia e secções do texto onde são apresentadas.

3.1. Definição e Harmonização da Nomenclatura

O primeiro passo no estudo feito será a definição da nomenclatura a usar no processo de classificação. Com imagens de satélite algumas classes de uso do solo poderão não ser identificáveis, como por exemplo, a distinção entre áreas comerciais e áreas industriais. Estas duas classes podem ser agrupadas como pertencentes a uma classe de cobertura mais geral que contém estas duas classes, denominada por territórios artificializados. Quando nos referimos a espaços destinados a agricultura, está-se a classificar o solo quanto ao uso que o ser humano lhe dá,

constituindo assim a classe de áreas agrícolas uma classe de uso. A partir de imagens de satélite de várias datas correspondentes a diferentes alturas do ano, espera-se que, com a alteração da resposta espectral, seja possível distinguir alguns tipos de áreas agrícolas de outras áreas com vegetação. Assim, definiu-se a seguinte nomenclatura de uso e cobertura do solo a partir do nível 1 da COS 2018:

1. *Territórios artificializados*: Maioritariamente constituída pela classe 1 da COS 2018. Nesta classe consideram-se áreas destinadas a atividades que envolvam a intervenção humana, como edifícios, áreas industriais e estradas. Exclui-se qualquer tipo de vegetação;
2. *Áreas agrícolas*: classe 2 da COS 2018, considerando como áreas utilizadas para agricultura incluindo pomares e olivais;
3. *Herbáceas*: Maioritariamente constituída pela classe 3 da COS 2018, considerando áreas ocupadas com vegetação essencialmente do tipo herbácea, incluindo jardins e parques urbanos presentes na classe 1 da COS 2018;
4. *Superfícies agroflorestais e florestais*: classes 4 e 5 da COS 2018. Nesta classe inclui-se parcelas com espécies florestais com ou sem culturas temporárias e/ou pastagens;
5. *Matos*: classe 6 da COS 2018, áreas naturais com vegetação espontânea, com coberto arbustivo superior ou igual a 25%;
6. *Espaços descobertos ou com pouca vegetação*: classe 7 da COS 2018, áreas naturais com pouca ou nenhuma vegetação, incluindo áreas com rocha nua, praias, areais e vegetação esparsa;
7. *Zonas húmidas*: classe 8 da COS 2018, áreas interiores ou litorais, cobertas temporariamente ou permanentemente por água;
8. *Corpos de água*: classe 9 da COS 2018, superfícies de água doce ou salgada.

Depois de definida a nomenclatura a usar no trabalho, procedeu-se à harmonização da nomenclatura, isto é, à correspondência das classes de cada um dos produtos a utilizar com a nomenclatura a ser usada. Como o nível 1 da maioria dos produtos é demasiado generalizado, é necessário recorrer a níveis mais pormenorizados. Para a COS, a partir da documentação oficial deste produto (Direção-Geral do Território, 2018, 2019) tem-se a correspondência de classes apresentada na Tabela 3.

Nomenclatura a adotar	COS 2015	COS 2018
1 Territórios artificializados	1 Territórios Artificializados, <u>excluindo:</u> 1.4.1 Espaços verdes urbanos e 1.4.2.01 Campos de golfe	1 Territórios artificializados, <u>excluindo:</u> 1.6.1.1 Campos de golfe e 1.7 Parques e jardins.
2 Áreas agrícolas	2 Áreas agrícolas e agroflorestais <u>excluindo:</u> 2.3 Pastagens permanentes e 2.4.4 Sistemas agroflorestais	2 Agricultura.
3 Herbáceas	1.4.1 Espaços verdes urbanos; 1.4.2.01 Campos de golfe; 2.3 Pastagens permanentes; 3.2.1 Vegetação herbácea natural.	1.6.1.1 Campos de golfe; 1.7 Parques e jardins; 3 Pastagens.
4 Superfícies agroflorestais e florestais	2.4.4 Sistemas agroflorestais; 3.1 Florestas.	4 Superfícies agroflorestais; 5 Florestas.
5 Matos	3.2.2 Matos.	6 Matos.
6 Espaços descobertos ou com pouca vegetação	3.3 Zonas descobertas e com pouca vegetação ou com vegetação esparsa.	7 Espaços descobertos ou com pouca vegetação.
7 Zonas húmidas	4 Zonas húmidas.	8 Zonas húmidas.
8 Corpos de água	5 Corpos de água.	9 Massas de água superficiais.

Tabela 3: Harmonização entre a nomenclatura a adotar e as classes da COS 2015 e as classes da COS 2018.

Relativamente à harmonização com o UA, como neste não existe distinção da classe dos matos, classe 5 na nomenclatura usada, decidiu-se agregar as classes 3 e 5 da nomenclatura usada. Desta forma, usando a documentação relativa ao UA 2012 (Copernicus, 2016), considerou-se a correspondência com a nomenclatura a usar apresentada na Tabela 4. Como se usa a COS 2018 para proceder à validação, tem-se também o processo de harmonização da nomenclatura apresentado na mesma tabela.

Nomenclatura a adotar	UA 2012	COS 2018
1 Territórios artificializados	1 Territórios artificializados, <u>excluindo</u> : 1.4.1 Áreas verdes urbanas.	1 Territórios artificializados, <u>excluindo</u> : 1.6.1.1 Campos de golfe e 1.7 Parques e jardins.
2 Áreas agrícolas	2 Áreas agrícolas, <u>excluindo</u> : 2.3 Pastagens.	2 Agricultura.
3 Herbáceas e Matos	1.4.1 Áreas verdes urbanas; 2.3 Pastagens; 3.2 Associações de vegetação herbácea.	1.6.1.1 Campos de golfe; 1.7 Parques e jardins; 3 Pastagens; 6 Matos.
4 Superfícies agroflorestais e florestais	3.1 Florestas	4 Superfícies agroflorestais; 5 Florestas.
5		
6 Espaços descobertos ou com pouca vegetação	3.3 Espaços abertos com pouca ou nenhuma vegetação.	7 Espaços descobertos ou com pouca vegetação.
7 Zonas húmidas	4 Zonas húmidas.	8 Zonas húmidas.
8 Corpos de água	5 Água.	9 Massas de água superficiais.

Tabela 4: Harmonização entre a nomenclatura a adotar e as classes do UA 2012 e as classes da COS 2018. Observação: a classe 5 deixa de existir e é agregada à classe 3.

O OSM2LULC é um software desenvolvido na Universidade de Coimbra para converter dados de OSM em MUCSs (Patriarca et al., 2019), resolvendo os problemas de:

- associação dos dados OSM a classes de ocupação e uso do solo: tem definida uma correspondência das *tags* do OSM com a nomenclatura de nível 2 da Corine Land Cover (CLC) e UA e do GlobeLand30;
- conversão de geometrias lineares do OSM em geometrias poligonais de área não nula: inclui ferramentas para a criação de *buffers* em torno das linhas de acordo com as características das zonas;
- inconsistências relacionados com a associação de classes diferentes às mesmas localizações: inclui a utilização de estratégias de análise geoespacial e topológicas para a seleção da classe a usar em cada situação.

Neste trabalho, utiliza-se uma versão modificada do software OSM2LULC para extrair um MUCS em formato vetorial seguindo a nomenclatura adotada. Na versão usada, o OSM2LULC associa *tags* do OSM às classes temáticas da CLC, cuja correspondência com a nomenclatura adotada é apresentada na Tabela 5. Nesta versão modificada do OSM2LULC, o processo de criação

do mapa vetorial termina antes de se resolver as inconsistências. Desta forma, ter-se-á a informação das várias classes que ocupam uma mesma área quando há sobreposição de classes.

Nomenclatura a adotar	OSM2LULC
1 Territórios artificializados	1.1 Zonas fabris; 1.2 Zonas industriais, comerciais e unidades de transporte; 1.3 Minas, aterros e zonas de construção; 1.4.2 Espaços de lazer e de desporto – excluindo campos de golfe.
2 Áreas agrícolas	2.1 Solo arável; 2.2 Culturas permanentes; 2.4 Áreas agrícolas heterogéneas.
3 Herbáceas	1.4.1 Espaços verdes urbanos; 2.3 Pastagens; 3.2.1 Pastagens naturais; 1.4.2 Espaços de lazer e de desporto – apenas campos de golfe.
4 Superfícies agroflorestais e florestais	3.1 Florestas.
5 Matos	3.2.4 Matos.
6 Espaços descobertos ou com pouca vegetação	3.3 Espaços abertos com pouca ou nenhuma vegetação.
7 Zonas húmidas	4 Zonas húmidas.
8 Corpos de água	5.1 Corpos de água no interior; 5.2 Água salgada.

Tabela 5: Harmonização entre a nomenclatura a adotar e a nomenclatura do CLC adotada no software OSM2LULC.

3.2. Criação de Grelhas

As grelhas a criar consistem num conjunto de polígonos quadrangulares que delimitam individualmente cada pixel da imagem. Cada polígono terá atributos a si associados em função da área que cada classe ocupa em cada um dos produtos usados, de acordo com a nomenclatura usada em cada um deles, nomeadamente a COS 2015, a COS 2018, o OSM2LULC e também o UA 2012, no caso da área de estudo A.

Dada a grande dimensão geográfica de cada área de estudo, este processo foi dividido na criação de várias grelhas. Para cada uma das áreas de estudo o processo é semelhante, sendo criadas grelhas de dimensão 10x10m correspondentes a janelas de 500x500 pixéis, tornando assim o processo menos dispendioso a nível computacional.

Numa primeira fase, são criadas as grelhas vazias, sem qualquer atributo associado, utilizando o código *Python* apresentado no Anexo A, recorrendo-se aos módulos de *osgeo* e *osr*.

Depois de criadas as grelhas vazias, é necessário associar a cada um dos polígonos a área ocupada por cada classe de cada um dos produtos. Para isso, é necessário começar por criar os atributos que guardarão a informação pretendida. De seguida, intersesta-se cada um dos polígonos da grelha com cada um dos polígonos existentes na COS 2015, COS 2018, OSM2LULC e UA 2012. Quando a interseção não for vazia, ao atributo correspondente à classe do produto a intersestar, na célula da grelha, atualiza-se a área ocupada pela respetiva classe/produto (ver Figura 6). No Anexo B apresenta-se a implementação deste processo para o caso da COS 2015, sendo o processo análogo para os restantes produtos.

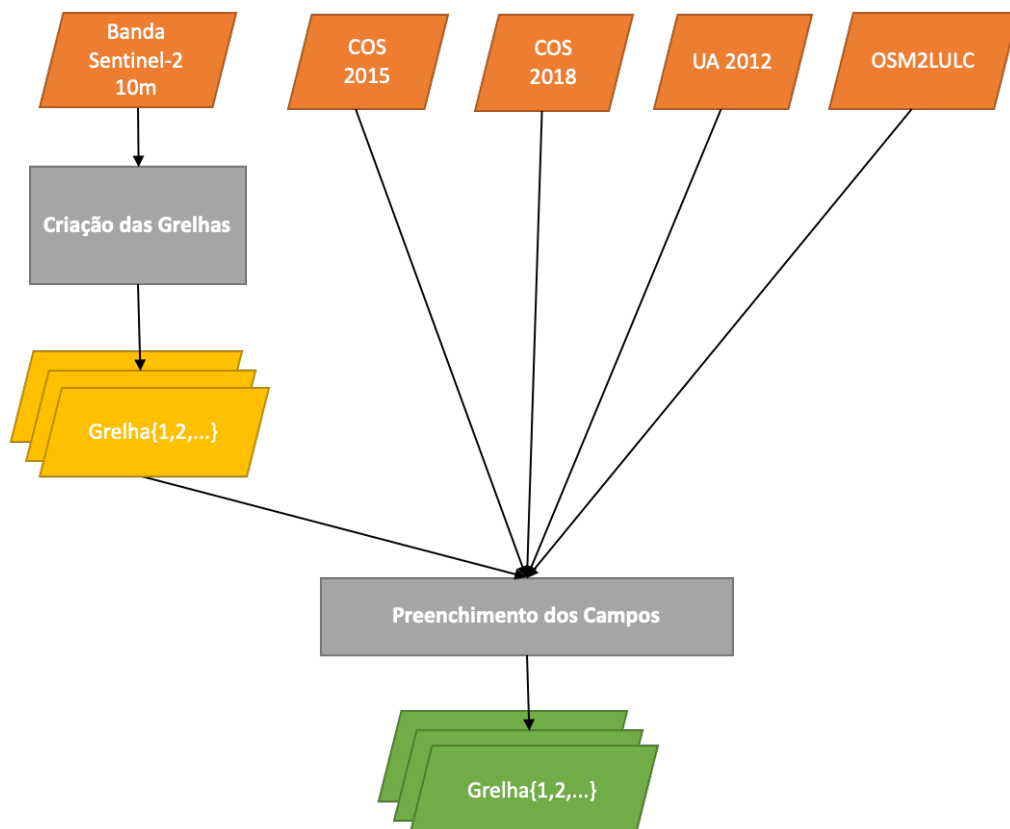


Figura 6: Fluxograma que ilustra o processo de criação das grelhas.

3.3. Criação de Dados de Treino

No processo de criação dos dados de treino recorre-se às grelhas criadas anteriormente, sendo criados para cada um dos produtos (COS 2015, OSM2LULC e UA 2012) áreas de treino às quais são associadas *features*, ou seja, variáveis que caracterizam cada área, sendo neste caso os valores dos pixels nas bandas do S2 e adicionalmente, nalguns casos, os índices radiométricos. Desta forma, a associação das áreas de treino com as *features* constituem os dados de treino.

Assim, criam-se áreas de treino exclusivamente a partir da COS 2015, exclusivamente do OSM, em simultâneo da COS 2015 e do OSM e exclusivamente do UA 2012. Às áreas de treino associam-se um conjunto de *features* de acordo com os critérios que serão explicados nesta secção de forma a constituírem os dados de treino.

3.3.1. Criação de Áreas de Treino

Para a criação das áreas de treino começa-se por recorrer à área ocupada por cada classe em cada um dos produtos. Esta área tomará um valor não negativo, sendo que uma área nula significará que a classe do produto não ocupa a região em questão e um valor positivo representa a área ocupada pela classe. Assim, para cada um dos produtos, extrai-se os polígonos a usar como treino para cada classe de acordo com os seguintes critérios de criação de áreas de treino:

0. Presença – consideram-se todos os polígonos da grelha que tenham associada à classe uma área positiva (denota-se por DT_*0 a área de treino criada a partir deste critério, onde ‘*’ tomará o código do produto usado para criar a área de treino);
1. Predominância – consideram-se todos os polígonos quadrangulares cuja classe ocupe maioritariamente essa área (denota-se por DT_*1 a área de treino criada a partir deste critério, onde ‘*’ tomará o código do produto usado para criar a área de treino);
2. Exclusividade – consideram-se todos os polígonos quadrangulares que tenham área de ocupação total (denota-se por DT_*2 a área de treino criada a partir deste critério, onde ‘*’ tomará o código do produto usado para criar a área de treino).

Cada área de treino é um conjunto de pixels sem variáveis (*features*) que os caracterizam. Desta forma, para cada área de estudo, são criados áreas de treino referentes a cada produto de

acordo com os critérios de criação de áreas de treino enunciados anteriormente. Assim, ter-se-á as áreas de treino:

- DT_C0, DT_C1 e DT_C2 extraídas exclusivamente da COS 2015 (Anexo C);
- DT_O0, DT_O1 e DT_O2 extraídas exclusivamente do OSM;
- DT_CO0, DT_CO1 e DT_CO2 extraídas em simultâneo da COS 2015 e do OSM;
- DT_U0, DT_U1 e DT_U2 extraídas exclusivamente do UA 2012.

Dada a grande extensão geográfica das áreas de treino que se retiram com os vários critérios, o esforço computacional é elevado. De forma a poder-se comparar a criação dos vários conjuntos de dados de treino, utiliza-se para classificação apenas uma amostra de cada área de treino (o código usado para a criação das amostras está disponível no Anexo D). Para fazer esta amostra, utilizam-se 10% das áreas de treino iniciais criadas de acordo com os seguintes critérios:

- S_DT_*2 – a amostra de DT_*2 é criada de forma a que, para cada classe, retire-se aproximadamente 10% dos pixéis;
- S_DT_*1 – a amostra de DT_*1 é criada de forma a que à amostra S_DT_*2, para cada classe, juntam-se cerca de 10% dos pixéis existentes no conjunto de pixéis existentes em DT_*1 e não existentes em DT_*2;
- S_DT_*0 – a amostra de DT_*0 é criada de forma a que, para cada classe, utiliza-se a amostra de pixéis S_DT_*1 da mesma classe e juntam-se os pixéis presentes noutras classes de S_DT_*1 que pertençam à área de treino DT_*0.

Para cada amostra a utilizar nas classificações, utilizando os critérios anteriores, tem-se assim para cada classe:

- Presença (código 0): pixéis exclusivos de DT_*2 na respetiva classe e pixéis exclusivos de DT_*1\DT_*2 que são usados como pertencentes a mais que uma classe;
- Predominância (código 1): pixéis exclusivos de DT_*2 e exclusivos de DT_*1\DT_*2 na respetiva classe;
- Exclusividade (código 2): pixéis exclusivos de DT_*2 na respetiva classe

3.3.2. Associação de *Features* às Áreas de Treino

Associados a cada um dos pixels, utiliza-se um conjunto de *features* obtidas das imagens de S2. Desta forma, tem-se os seguintes critérios de seleção de *features* para cada pixel:

- F1 – apenas as bandas de 10m (2, 3, 4 e 8) para as datas selecionadas;
- F1-4 –bandas de 10m (2, 3, 4 e 8) para as datas selecionadas no pixel e nos 4 pixels vizinhos (Norte, Sul, Este e Oeste);
- F2 – todas as bandas convertidas em bandas com resolução espacial de 10m para as datas selecionadas;
- F3 - todas as bandas convertidas em bandas com resolução espacial de 10m e índices radiométricos NDVI, NDWI, NDBI e SAVI para as datas selecionadas.

A associação das *features* às áreas de treino constituem os dados de treino, cujo processo de criação (secção 3.3.1 e secção 3.3.2) se encontra resumido no fluxograma da Figura 7.

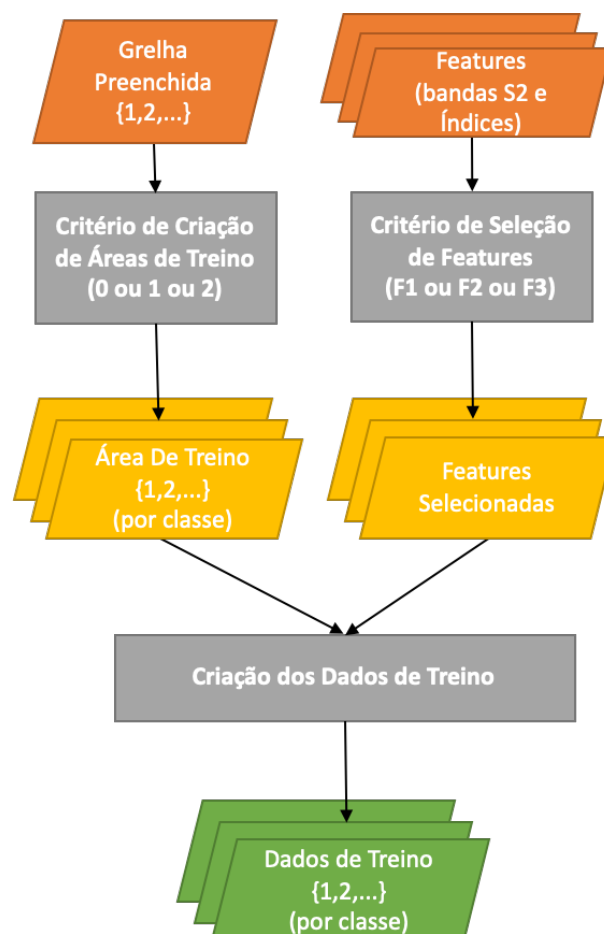


Figura 7: Fluxograma que ilustra o processo de criação de dados de treino (sem filtragem com índices radiométricos).

3.3.3. Processo de Filtragem das Áreas de Treino com Índices Radiométricos

Dado que os índices radiométricos constituem uma fonte de informação adicional para cada pixel, espera-se que o seu comportamento seja semelhante em cada uma das classes. Estes índices podem assim servir como uma fonte de informação adicional que caracteriza o pixel, ou seja, podem ser usados como *features* adicionais a associar aos dados de treino (no caso F3) e/ou serem utilizados para filtrar as áreas de treino. Para usar estes índices como fonte de informação para filtrar as áreas de treino, utilizam-se as abordagens descritas nos parágrafos seguintes.

Uma primeira forma de utilizar os índices radiométricos para filtrar as áreas de treino envolve a identificação de valores a partir da literatura e a criação e análise de histogramas para cada índice por classe.

Uma segunda hipótese de filtrar as áreas de treino utilizando os índices radiométricos, será recorrendo à distribuição dos valores dos índices para cada uma das datas das imagens usadas. Utilizando os índices NDVI e NDWI, para cada uma das datas e para cada classe, determina-se a média (m) dos valores e o desvio padrão (d). Com esta informação estatística, os dados com valores fora dos intervalos $]m-3d, m+3d[$ são excluídos dos dados de treino. Desta forma assume-se que os dados seguem uma lei Normal de parâmetros média (m) e desvio padrão (d) para cada uma das variáveis (NDVI e NDWI).

Assim, definiram-se duas formas possíveis de se realizarem filtrações a partir dos índices radiométricos:

- Critério L: utilizando condições pré-estabelecidas;
- Critério M: utilizando a informação estatística: média e desvio padrão de cada índice.

O processo de filtragem das áreas de treino a partir dos índices radiométricos é assim esquematizado na Figura 8.

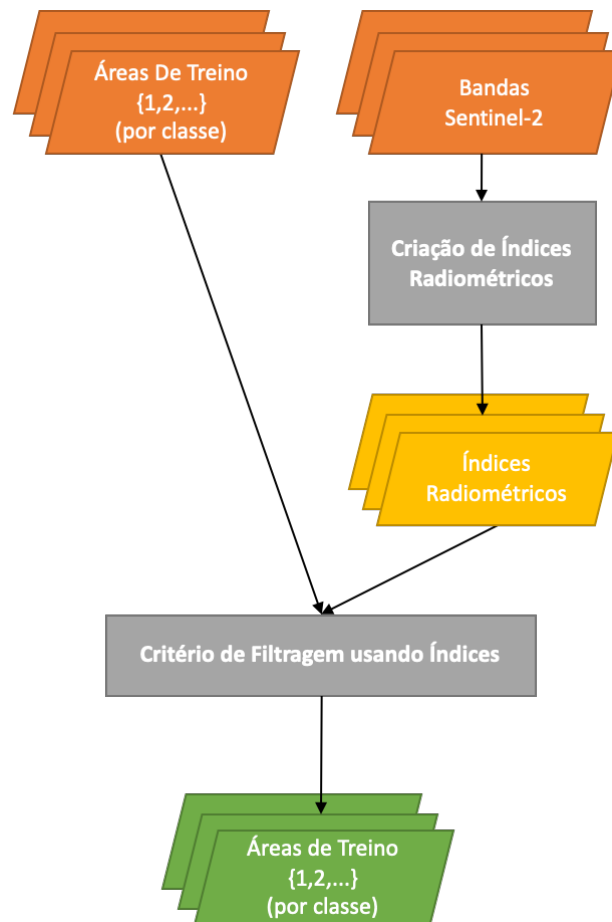


Figura 8: Fluxograma que ilustra o processo de criação de áreas de treino com filtragem usando índices radiométricos.

3.3.4. Denominação das Áreas de Treino e Dados de Treino

Dada a grande quantidade de dados de treino que se podem criar a partir dos vários produtos, na Tabela 6 são apresentadas as denominações dos dados de treino, e das respetivas áreas de treino, que são criados com a metodologia descrita acima. Quando se pretende referir apenas às áreas de treino, isto é, aos dados de treino sem nenhuma *feature* associada, a denominação é semelhante sem conter o código final “_F1” ou “_F2” ou “_F3”.

Denominação dos Dados de Treino	Denominação das Áreas de Treino	Fonte de dados	Critério	Features	Filtragem com Índices Radiométricos
DT_C0_F1	DT_C0	COS 2015	0 (Presença)	F1	-
DT_C0_F2	DT_C0	COS 2015	0 (Presença)	F2	-
DT_C0_F3	DT_C0	COS 2015	0 (Presença)	F3	-
DT_C1_F1	DT_C0	COS 2015	1 (Predominância)	F1	-
DT_C1_F1-4	DT_C1	COS 2015	1 (Predominância)	F1-4	-
DT_C1_F2	DT_C1	COS 2015	1 (Predominância)	F2	-
DT_C1L_F2	DT_C1L	COS 2015	1 (Predominância)	F2	Critério L
DT_C1M_F2	DT_C1M	COS 2015	1 (Predominância)	F2	Critério M
DT_C1_F3	DT_C1	COS 2015	1 (Predominância)	F3	-
DT_C2_F1	DT_C2	COS 2015	2 (Exclusividade)	F1	-
DT_C2_F2	DT_C2	COS 2015	2 (Exclusividade)	F2	-
DT_C2_F3	DT_C2	COS 2015	2 (Exclusividade)	F3	-
DT_O0_F2	DT_O0	OSM	0 (Presença)	F2	-
DT_O1_F2	DT_O1	OSM	1 (Predominância)	F2	-
DT_O2_F2	DT_O2	OSM	2 (Exclusividade)	F2	-
DT_CO0_F2	DT_CO0	COS 2015 &	0 (Presença)	F2	-
DT_CO1_F2	DT_CO1	COS 2015 &	1 (Predominância)	F2	-
DT_CO2_F2	DT_CO2	COS 2015 &	2 (Exclusividade)	F2	-
DT_U0_F2	DT_U0	UA 2012	0 (Presença)	F2	-
DT_U1_F2	DT_U1	UA 2012	1 (Predominância)	F2	-
DT_U2_F2	DT_U2	UA 2012	2 (Exclusividade)	F2	-

Tabela 6: Denominações dos dados de treino consoante os critérios e *features* utilizados.

3.4. Avaliação dos Dados de Treino

Para avaliar os dados de treino são utilizadas duas abordagens: 1) comparar as áreas de treino com o MUCS de referência; 2) avaliar a separabilidade dos dados de treino. A criação do MUCS de referência é feita a partir da COS 2018, utilizando o critério 1 (Predominância) de forma a converter para formato *raster* a COS 2018 que é disponibilizada em formato vetorial. Desta forma, todos os pixels na área de estudo estão classificados com uma e uma só classe. Considerando 2 *clusters*, um constituído pelos dados da classe a avaliar e outro constituído pelos dados das classes complementares, são calculadas inicialmente as matrizes de covariância e as médias para cada *cluster*. Estes parâmetros estatísticos são depois utilizados no cálculo da distância de Bhattacharya de forma a avaliar a separabilidade dos dados. Uma das principais características desta distância é

o facto de esta ter em conta a forma como os dados se distribuem no espaço espectral, isto é, tem em conta a dispersão dos dados, cuja fórmula de cálculo (9) é a seguinte:

$$BD(i, i^c) = \frac{1}{8} (\mu_i - \mu_{i^c})^T \left[\frac{\Sigma_i + \Sigma_{i^c}}{2} \right]^{-1} (\mu_i - \mu_{i^c}) + \frac{1}{2} \ln \left(\frac{|\frac{\Sigma_i + \Sigma_{i^c}}{2}|}{(|\Sigma_i| |\Sigma_{i^c}|)^{\frac{1}{2}}} \right) \quad (9)$$

com

i conjunto de dados da classe i (núcleo i);

i^c conjunto de dados pertencentes todas as classes exceto à classe i (núcleo i^c);

$BD(i, i^c)$ distância de Bhattacharya entre o núcleo i e o núcleo i^c ;

μ_i, μ_{i^c} médias dos núcleos i e i^c , respetivamente;

Σ_i, Σ_{i^c} matrizes de covariância dos núcleos i e i^c , respetivamente.

É de notar que a separabilidade dos dados depende das *features* a utilizar, obtendo-se assim valores diferentes quando são usadas 4 bandas, todas as bandas e todas as bandas em conjunto com os índices radiométricos.

A comparação das áreas de treino com o MUCS é feita calculando a percentagem de pixéis onde há concordância da classe atribuída nas áreas de treino com o MUCS de referência, usando a metodologia explicada em “3.6. Validação”.

3.5. Classificação e Generalização

O processo de classificação é feito com o classificador *Random Forest*, implementado no módulo *scikit-learn* de *Python*, utilizando os dados de treino gerados de acordo com os enumerados na Tabela 6.

Por motivos de esforço computacional, utiliza-se as definições padrão do classificador. Depois da classificação, dado que se pretende obter MUCSs semelhantes à COS 2018, é aplicado um filtro com uma janela com o formato de disco (com um raio a definir) e o pixel é reclassificado para a classe com maior ocupação na janela. A este processo denominemos por generalização do MUCS. Este é feito recorrendo ao módulo de *Python skimage*, utilizando um disco com um raio de 4 pixéis. Na Figura 9 são apresentadas as etapas do processo de classificação e generalização para a produção de MUCS.

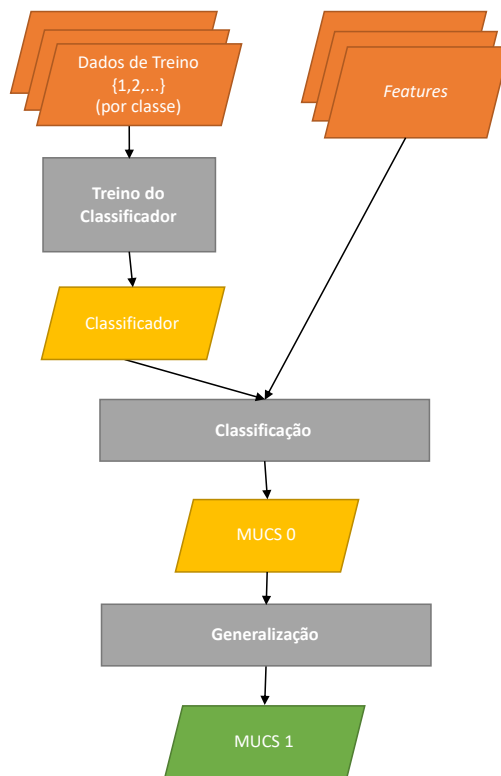


Figura 9: Fluxograma que ilustra o processo de classificação e sua generalização.

3.6. Validação

Para o processo de validação (Figura 10) é utilizado o MUCS de referência criado a partir da COS 2018 sendo considerado para cada pixel a classe majoritária, ou seja utilizando o critério 2 (Predominância). Este mapa de referência é utilizado para validar os dados de treino, as amostras criadas a partir deles, bem como os MUCSs resultantes da classificação e da generalização. Assim, os dados de treino criados e os MUCSs obtidos das classificações e das generalizações são comparados com o MUCS de referência, construindo matrizes de confusão para cada mapa a avaliar e calculando a exatidão global e os índices de exatidão do utilizador e do produtor para cada classe (os mapas de referência obtidos estão disponíveis no Anexo E).

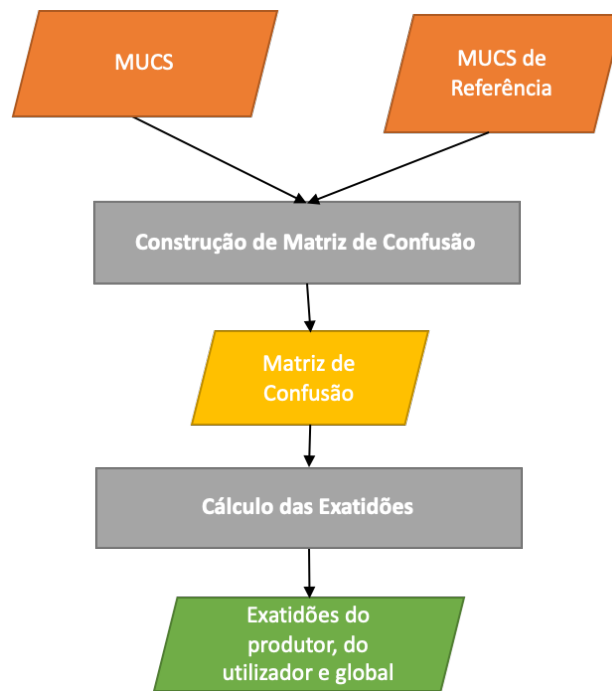


Figura 10: Fluxograma que ilustra o processo de validação dos MUCSs criados.

No caso da COS, utilizando a COS 2015 para obter dados de treino e a COS 2018 para obter o mapa de referência, pode-se determinar as áreas onde ocorreram mudança de uso e ocupação de solo. Desta forma, para complementar a avaliação de um MUCS pode-se calcular a proporção de pixels corretamente classificados nestas áreas.

4. Casos de Estudo

Definindo duas áreas de estudo em Portugal Continental, aplicou-se a metodologia de criação de dados de treino para produção de MUCSs apresentada no capítulo anterior. Assim, começa-se por apresentar na secção 4.1 as duas áreas de estudo escolhidas (A e B) e os produtos a usar para cada uma das regiões. Na secção seguinte apresentam-se os resultados obtidos aplicando as metodologias de criação de dados de treino apresentadas anteriormente.

4.1. Áreas de Estudo

Foram escolhidas duas áreas de estudo em Portugal Continental, uma situada na zona de Lisboa e outra na Serra da Estrela. Para cada uma das zonas, seleccionou-se um conjunto de 3 imagens de S2 com datas distribuídas ao longo do ano de 2018, optando-se assim por se realizar classificações utilizando séries temporais de imagens correspondentes a diferentes estações do ano. Desta forma, com os vários estágios da vegetação ao longo do ano espera-se que as classes com vegetação sejam mais facilmente distinguíveis.

Nas subsecções seguintes descrevem-se as áreas de estudo, apresentando as imagens de S2 e os produtos disponíveis utilizados.

4.1.1. Área de Estudo A: Lisboa

A primeira zona escolhida, contendo a cidade de Lisboa, é apresentada e enquadrada em Portugal Continental na Figura 11 a). Esta área de estudo ocupa uma área de aproximadamente 1560km² e situa-se na zona centro-sul de Portugal Continental. Esta região é maioritariamente do tipo urbano, situando-se numa área costeira de Portugal, incluindo assim uma grande parte de água

do Oceano Atlântico. Na Figura 11, em b), c) e d), são apresentadas composições de cor natural das imagens de S-2 a utilizar, apresentando-se na Tabela 7 os detalhes das mesmas.

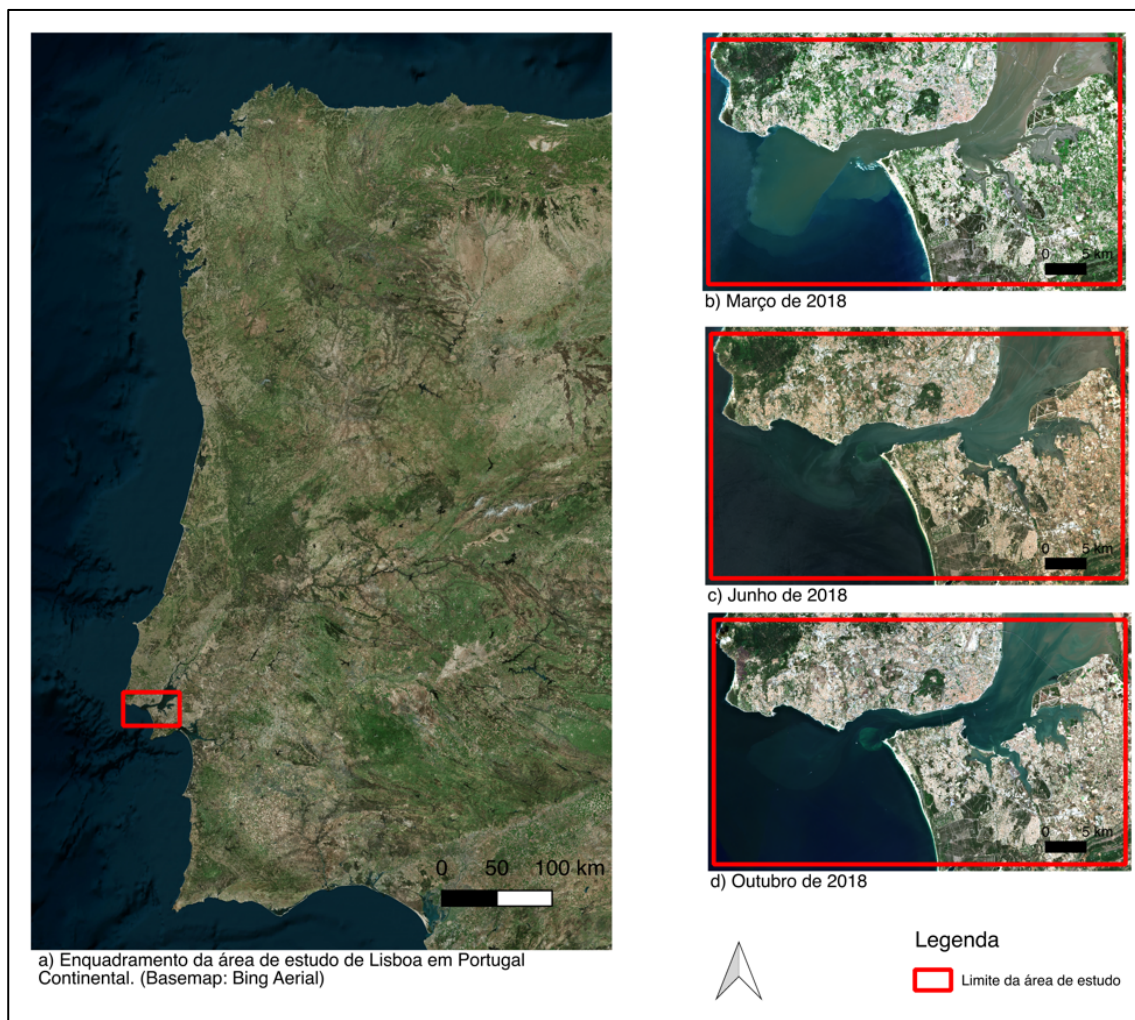


Figura 11: Área de estudo A – a) enquadramento da área de estudo em Portugal Continental (Basemap: Bing Aerial); composições a cor natural, RGB (bandas 4-3-2), das imagens de Sentinel-2 para as datas referentes aos meses de março (b), de junho (c) e outubro (d) de 2018.

Imagem da Figura 11	Satélite	Tipo de Produto	Data de Aquisição	Grelha Sentinel
b)	Sentinel-2A	Nível-2A	21/março/2018	T29SMC
c)	Sentinel-2A	Nível-2A	19/junho/2018	T29SMC
d)	Sentinel-2B	Nível-2A	22/outubro/2018	T29SMC

Tabela 7: Detalhes das imagens de Sentinel-2 utilizadas, para a área de estudo A.

Nas figuras 12, 13, 14 e 15 são apresentadas para a área de estudo A a COS 2015, a COS 2018, o OSM e o UA 2012, respectivamente. Observa-se que esta zona é diversificada quanto ao uso e cobertura do solo, incluindo áreas artificializadas, áreas agrícolas, áreas com vegetação, zonas húmidas e corpos de água. No entanto, esta área é predominantemente composta por água e zonas artificializadas.

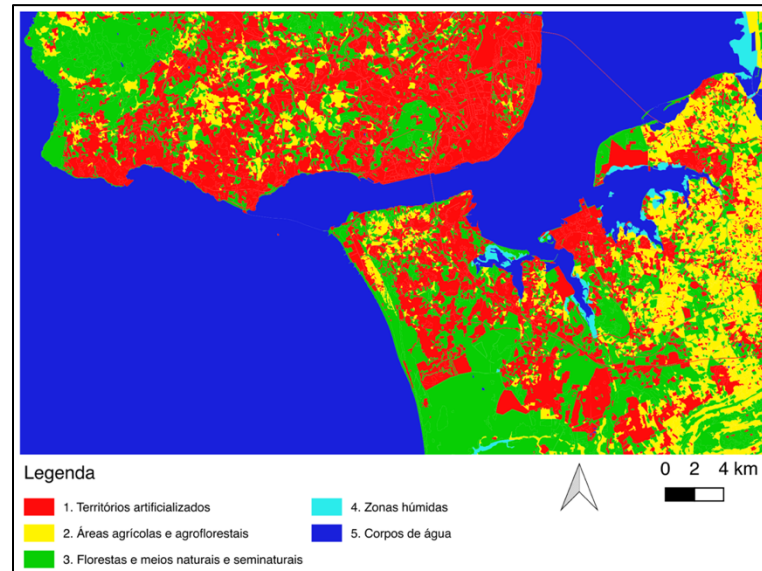


Figura 12: Área de estudo A – Nível 1 da COS 2015.

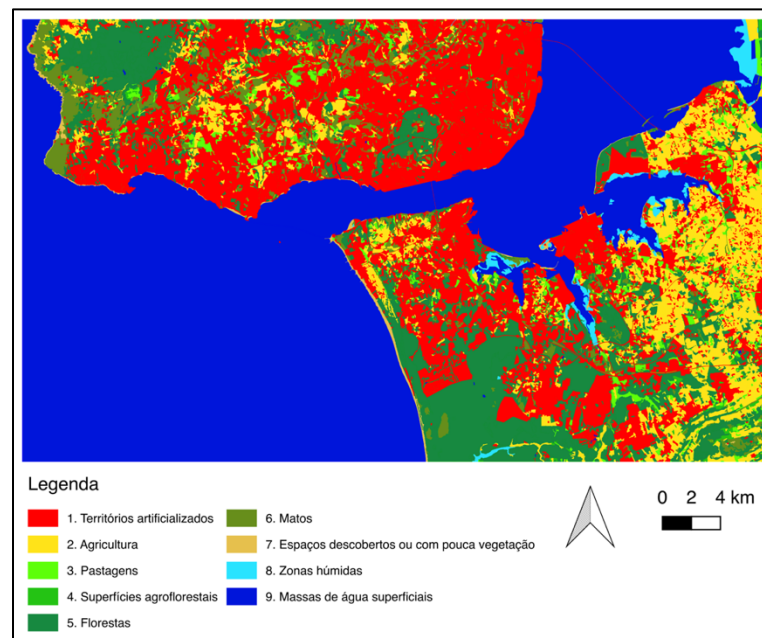


Figura 13: Área de estudo A – Nível 1 da COS 2018.

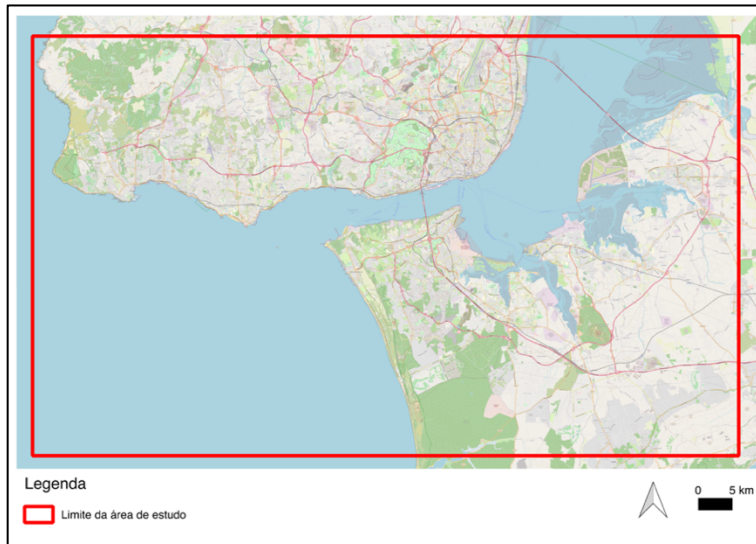


Figura 14: Área de estudo A – OpenStreetMap (Basemap OpenStreetMap no QGIS).

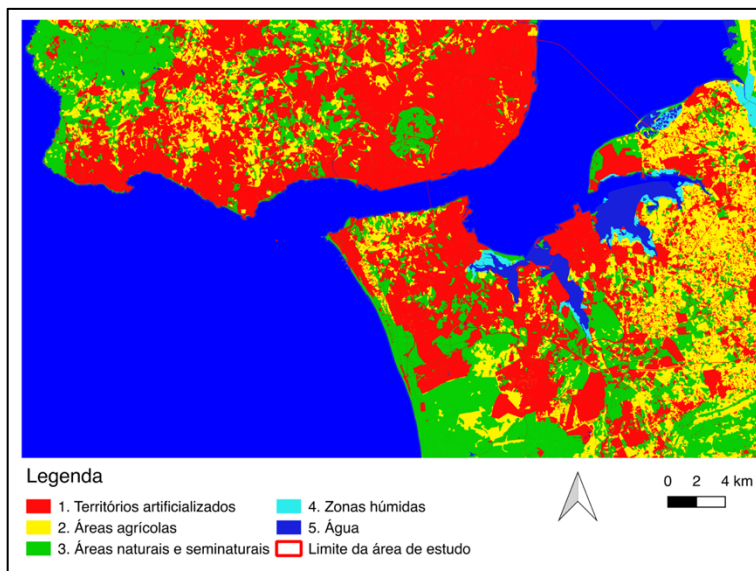


Figura 15: Área de estudo A – Nível 1 do UA 2012.

4.1.2. Área de Estudo B: Serra da Estrela

A segunda zona escolhida contém o Parque Nacional da Serra da Estrela. Na Figura 16 a) é apresentado o seu enquadramento em Portugal Continental. Esta área de estudo ocupa uma área de aproximadamente 2000km² e situa-se na zona centro-norte de Portugal Continental. Esta região é maioritariamente do tipo rural, situando-se numa área interior de Portugal. Na Figura 16, em b), c)

e d), são apresentadas composições de cor natural das imagens de S2 a utilizar, cujos detalhes são apresentado na Tabela 8.

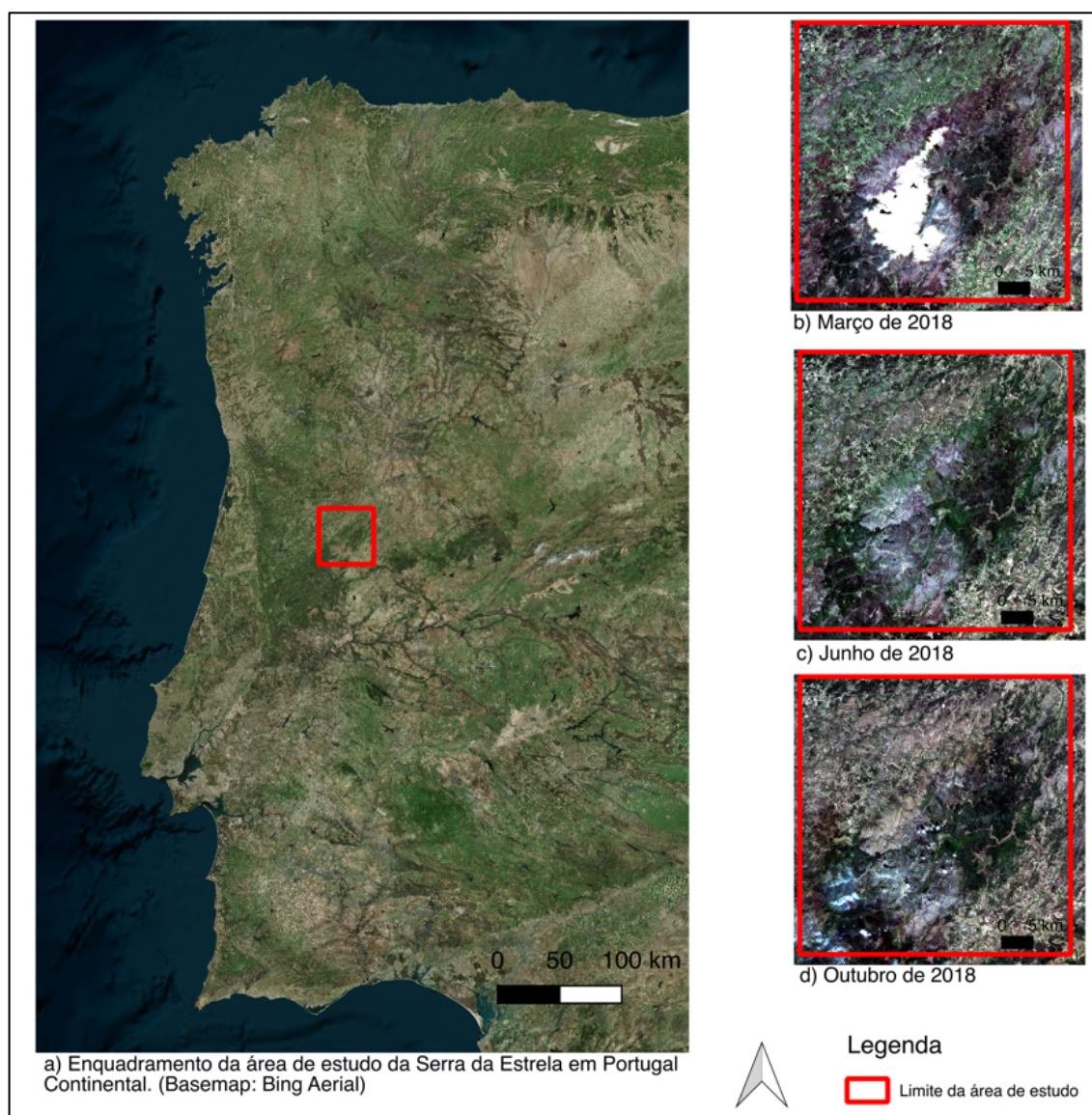


Figura 16: Área de estudo B – a) enquadramento da área de estudo em Portugal Continental (Basemap: Bing Aerial); composições a cor natural, RGB (bandas 4-3-2), das imagens de Sentinel-2 para as datas: b) Março de 2018; c) Junho de 2018; d) Outubro de 2018.

Imagem da Figura	Satélite	Tipo de Produto	Data de Aquisição	Grelha Sentinel
16				
b)	Sentinel-2B	Nível-2A	26/março/2018	T29TPE
c)	Sentinel-2A	Nível-2A	19/junho/2018	T29TPE
d)	Sentinel-2B	Nível-2A	22/outubro/2018	T29TPE

Tabela 8: Detalhes das imagens de Sentinel-2 utilizadas para a área de estudo B.

Nas figuras 17, 18 e 19 são apresentadas para a área de estudo B a COS 2015, a COS 2018 e o OSM, respetivamente. Observa-se que esta zona é diversificada quanto ao uso e cobertura do solo, sendo predominantemente ocupada por classes de vegetação, incluindo matos e floresta, não contendo zonas húmidas. No caso do OSM é de notar que existe uma grande área em torno do Parque Natural da Serra da Estrela sem dados.

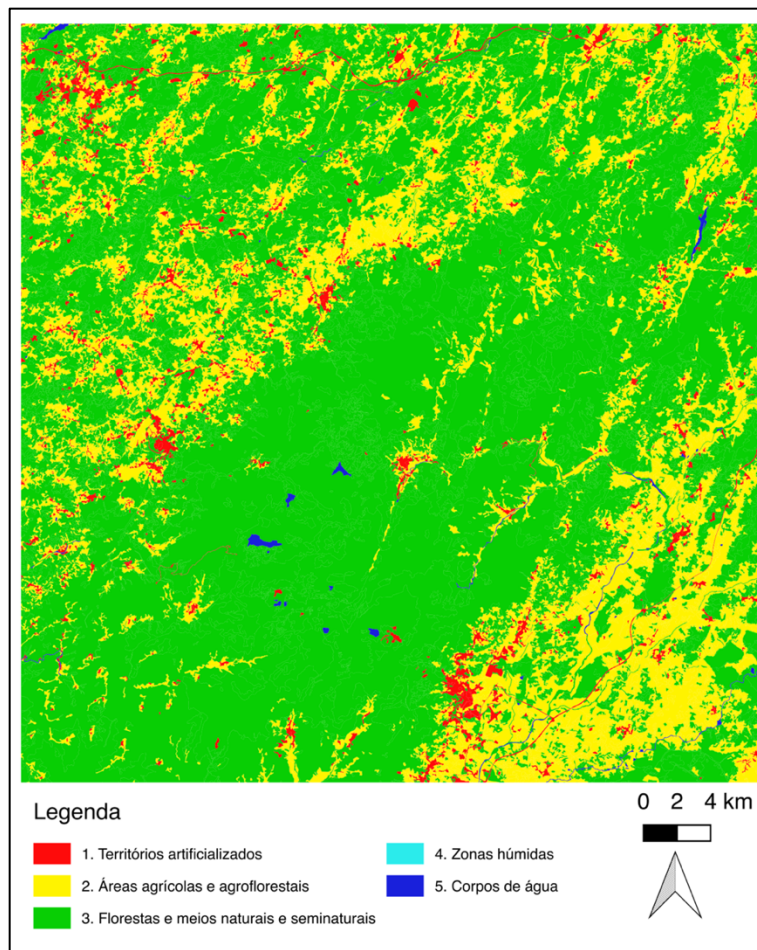


Figura 17: Área de estudo B – Nível 1 da COS 2015.

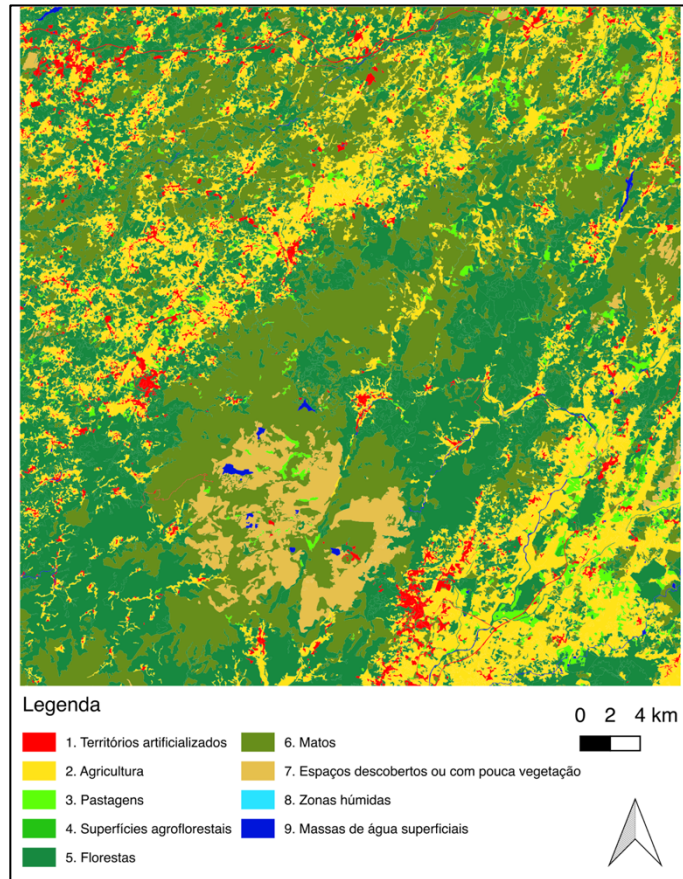


Figura 18: Área de estudo B – Nível 1 da COS 2018.

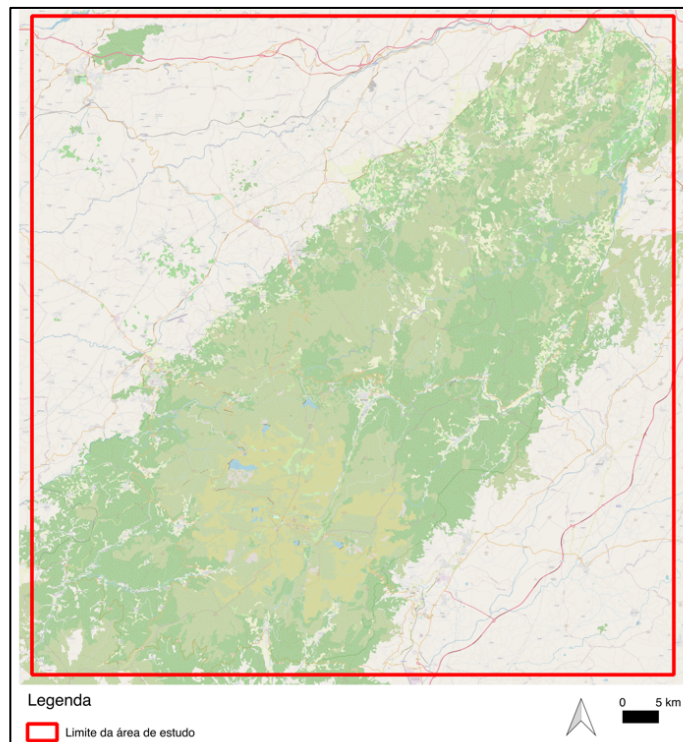


Figura 19: Área de estudo B – OpenStreetMap (Basemap QGIS: OpenStreetMap).

4.2. Resultados e Discussão

Nesta secção apresentam-se os resultados obtidos nestes casos de estudo. Inicialmente, em 4.2.1, apresentam-se os resultados obtidos com as classificações feitas com os dados de treino extraídos diretamente da COS 2015. Na subsecção seguinte, 4.2.2, mostra-se as classificações feitas com os dados de treino anteriores com filtragens a partir dos índices radiométricos. Na subsecção 4.2.3 expõem-se as classificações obtidas com os dados de treino extraídos do OSM. Em seguida, em 4.2.4, utilizam-se em simultâneo os dados de treino extraídos diretamente da COS 2015 e do OSM. Depois, em 4.2.5, fazem-se classificações com os dados de treino extraídos diretamente do UA 2012. Por fim, na última subsecção, 4.2.6, utilizam-se dados de treino extraídos diretamente da COS 2015 mas usando como *features* um conjunto de variáveis que caracterizam o próprio pixel e os seus vizinhos.

4.2.1. Classificação com dados de treino extraídos da COS 2015

Começou-se por usar a COS 2015 para criar os primeiros conjuntos de dados de treino. As Figura 20 e Figura 21 mostram os dados usados para treino, observando-se que a área total coberta pelas áreas de treino engloba toda a área de estudo quando estas são criadas a partir dos critérios Presença (C0) ou Predominância (C1). A COS 2015 ocupa toda a extensão dos limites das áreas de estudo, estando toda a área interior classificada, não havendo assim áreas que não entrem nestes dados de treino. As áreas que são eliminadas de DT_C1 para DT_C2 correspondem a áreas que são fronteira dos polígonos da COS 2015. Desta forma, um pixel que contenha a fronteira entre polígonos conterá duas ou mais classes, ou seja, não terá uma ocupação total de uma só classe como quando usado o critério Exclusividade (C2).

Assim, como se pode ver na Figura 20 e na Figura 21, tem-se que a área coberta pelas áreas de treino DT_C0 e DT_C1 é constituída pelas áreas a amarelo em conjunto com as áreas a azul. No caso da área coberta pelas áreas de treino DT_C2 esta é apresentada unicamente a amarelo. Note-se que a área coberta por DT_C0 coincide com a área coberta por DT_C1, e que a área coberta por DT_C2 está contida em DT_C0. No caso de DT_C0 os pixéis a azul entram para todas as classes enquanto que para DT_C1 cada pixel entra para a classe maioritária.

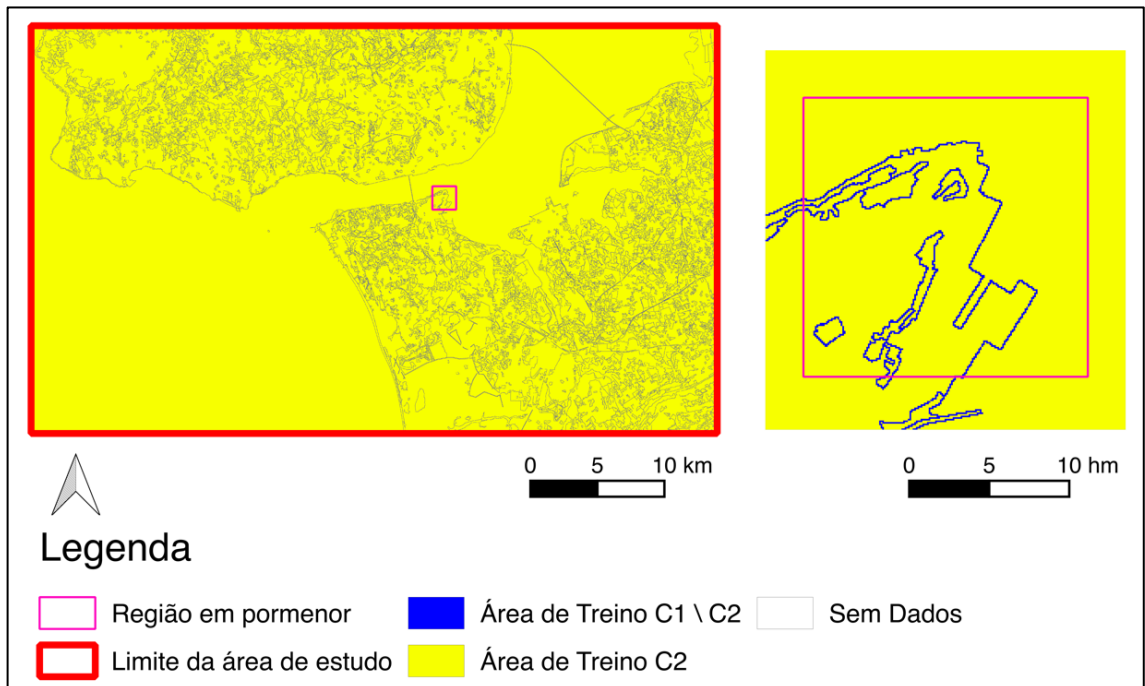


Figura 20: Área de estudo A – Áreas cobertas pelos dados de treino obtidos a partir da COS 2015 (azul – área coberta por DT_C0 e DT_C1; amarelo – área pertencente a DT_C0, DT_C1 e DT_C2).

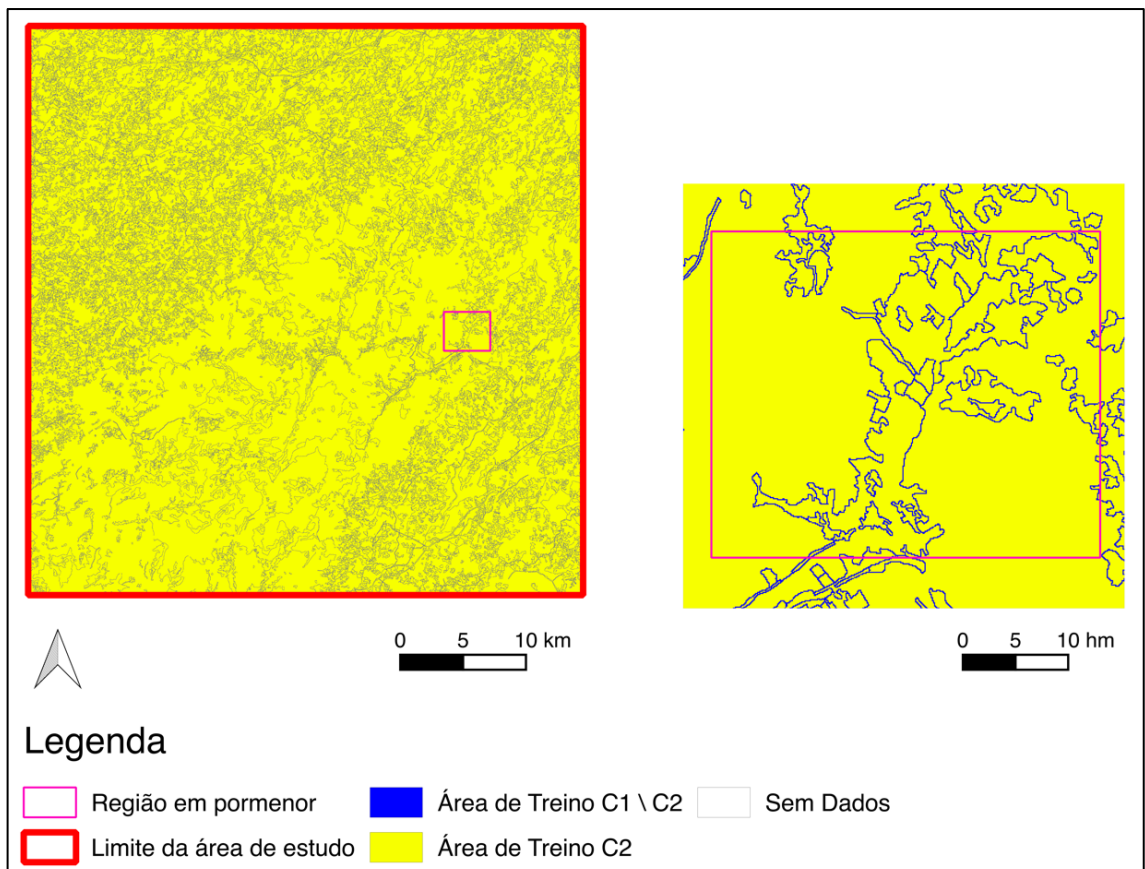


Figura 21: Área de estudo B – Áreas cobertas pelos dados de treino obtidos a partir da COS 2015 (azul – área coberta por DT_C0 e DT_C1; amarelo – área pertencente a DT_C0, DT_C1 e DT_C2).

Na Tabela 9 e na Tabela 10 são apresentadas as comparações dos dados de treino criados para a área de estudo A e para a área de estudo B, respetivamente, com os respetivos mapas de referência de cada área. Observa-se que, para ambas as áreas de estudo, a qualidade dos dados de treino é melhor usando o critério C2 (Exclusividade). A qualidade destes dados de treino é muito semelhante quando se usa o critério C1 (Predominância) que por sua vez é superior aos criados utilizando o critério C0 (Presença). Para ambas as áreas, verifica-se que a qualidade dos dados de treino para a classe 3 (Herbáceas) é muito inferior às restantes classes, indicando que possivelmente em ambas as áreas de estudo existam áreas que estavam mal classificadas em 2015 ou que a sua ocupação tenha mudado de 2015 para 2018. De forma global, os dados de treino obtidos pela COS 2015 utilizando os critérios C0, C1 e C2 apresentam uma exatidão superior a 85%.

A partir de ambas as tabelas, tem-se que a qualidade dos dados de treino quer usando a sua totalidade quer usando uma amostra destes (denotadas por S_DT_C0, S_DT_C1 e S_DT_C2) são muito semelhantes, nunca diferindo em módulo mais de 0,5%. Assim, conclui-se que as amostras criadas em termos de exatidão são muito semelhantes ao conjunto de dados de treino completos.

	DT_C0		S_DT_C0		DT_C1		S_DT_C1		DT_C2		S_DT_C2	
	Nº Pixéis	%	Nº Pixéis	%	Nº Pixéis	%	Nº Pixéis	%	Nº Pixéis	%	Nº Pixéis	%
Classe 1	3972355	93,2	397166	93,2	3739448	98,7	373943	98,7	3507439	98,7	350743	98,7
Classe 2	1437581	88,2	143771	88,3	1321713	95,3	132170	95,4	1207928	95,5	120792	95,5
Classe 3	777523	55,6	77889	55,6	683898	63,1	68389	63,2	593924	63,3	59392	63,3
Classe 4	2270159	90,3	226863	90,4	2130717	95,9	213071	95,9	1994581	96,2	199458	96,2
Classe 5	714582	79,7	71479	79,7	638459	88,9	63845	88,9	565285	89,1	56528	89,2
Classe 6	82062	72,7	8217	73,1	68600	86,9	6859	87,5	55515	89,6	5551	89,8
Classe 7	132567	93,2	13245	93,3	123733	99,8	12372	99,8	115058	99,9	11505	99,9
Classe 8	6909611	99,7	690946	99,7	6888952	100,0	688895	100,0	6868692	100,0	686869	100,0
Global	16296440	92,6	1629576	92,6	15595520	96,6	1559544	96,6	14908422	96,9	1490838	96,9

Tabela 9: Área de estudo A – Validação dos dados de treino criados a partir da COS 2015 e respetivas amostras: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos para DT_C0, DT_C1 e DT_C2, e para as amostras deles extraídas, denotadas, respetivamente, por S_DT_C0, S_DT_C1 e S_DT_C2.

	DT_C0		S_DT_C0		DT_C1		S_DT_C1		DT_C2		S_DT_C2	
	Nº Pixéis	%	Nº Pixéis	%	Nº Pixéis	%	Nº Pixéis	%	Nº Pixéis	%	Nº Pixéis	%
Classe 1	739242	82,9	73859	82,9	626710	97,4	62670	97,4	517552	97,4	51755	97,4
Classe 2	4733780	87,4	473464	87,4	4312403	95,2	431240	95,2	3896943	95,4	389694	95,4
Classe 3	1042680	36,2	104303	36,1	894723	42,0	89471	42,0	753668	42,9	75366	42,8
Classe 4	8220968	90,6	821988	90,6	7730477	95,9	773047	95,9	7247545	95,9	724754	95,9
Classe 5	5911309	89,5	591098	89,5	5588060	94,4	558805	94,4	5270267	94,7	527026	94,7
Classe 6	826966	93,7	82714	93,8	785958	98,5	78595	98,5	745296	98,9	74529	99,0
Classe 7	0	-	0	-	0	-	0	-	0	-	0	-
Classe 8	68348	83,0	6858	82,8	57089	99,3	5708	99,4	46161	99,3	4616	99,6
Global	21543293	86,8	2154284	86,8	19995420	93,1	1999536	93,1	18477432	93,5	1847740	93,5

Tabela 10: Área de estudo B – Validação dos dados de treino criados a partir da COS 2015 e respectivas amostras: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos para DT_C0, DT_C1 e DT_C2, e para as amostras deles extraídas, denotadas, respetivamente, por S_DT_C0, S_DT_C1 e S_DT_C2.

Os gráficos na Figura 22 e na Figura 23, relativos às áreas de estudo A e B, respetivamente, mostram a separabilidade das classes para cada uma das amostras e conjunto de *features* considerado conforme é explica na secção 3.4. A leitura do gráfico, em função do conjunto de *features*, deve ser lida tendo em conta o pico de cada barra. Por exemplo, na área de estudo A, para a classe 7 com o critério C2, usando as *features* F1 é de aproximadamente 2,3, usando F2 é aproximadamente 4,8 e usando F3 é aproximadamente 10,4. Pode-se observar que, para ambos os casos, em cada classe a separabilidade é maior quanto maior o número de *features* a usar, isto é, quanto maior o número de bandas utilizadas para caracterizar um pixel maior é a separabilidade. Este resultado poderá ser explicado pelo aumento da dimensionalidade do problema, significando que quando se aumenta o número de características que caracterizam um padrão, este é mais fácil de ser distinguido dos outros.

Verifica-se também que para cada classe, a separabilidade é maior quando se usam pixéis puros (critério C2 – Exclusividade). Usando o critério C0 (Presença) a separabilidade é menor que usando o critério C1 (Predominância) que por sua vez é menor que usando o critério C2 (Exclusividade). Na primeira iteração, de DT_C0 para DT_C1, excluem-se pixéis que pertencem a mais que uma classe, obrigando estes a pertencer a uma só classe (classe maioritária). Desta forma, um pixel que pertence a duas classes, em DT_C0 pertencerá a dois núcleos de dados, ao da classe em análise e ao das classes complementares, havendo assim uma menor distância entre os dois núcleos, ou seja, uma menor separabilidade. De DT_C1 para DT_C2, ao excluírem-se os pixéis mistos, excluem-se pixéis com respostas espectrais próximas e que pertencem a classes diferentes. Desta forma, aumenta-se a distância entre os dois núcleos de dados, ou seja, aumenta-se a separabilidade de uma classe em relação aos dados das restantes classes.

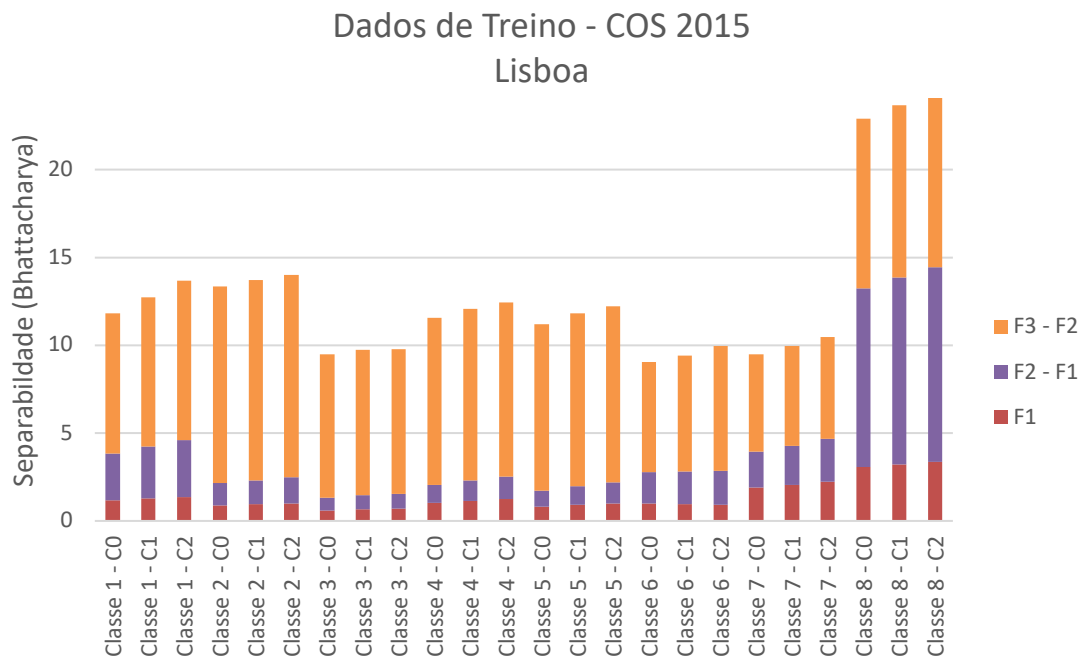


Figura 22: Área de estudo A – Avaliação da separabilidade da amostra de dados de treino obtidos diretamente pela COS 2015 usando a distância de Bhattacharya.

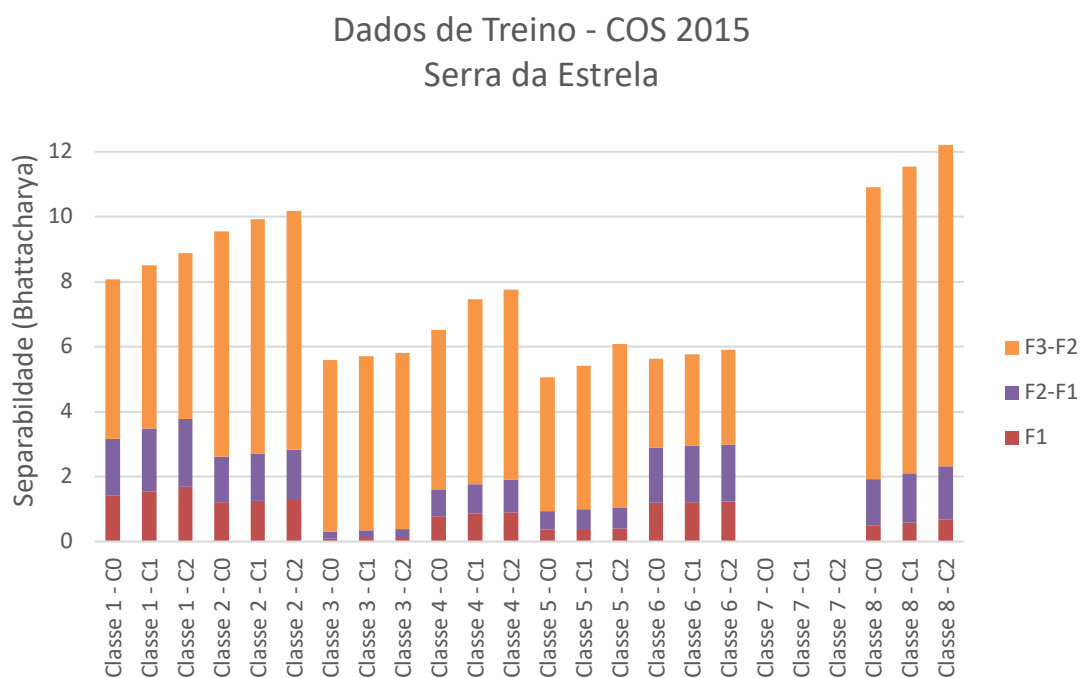


Figura 23: Área de estudo B – Avaliação da separabilidade da amostra de dados de treino obtidos diretamente pela COS 2015 usando a distância de Bhattacharya.

Depois de criadas as áreas de treino, procede-se à classificação das imagens de satélite e à sua generalização.

Nas tabelas 11 e 13 apresentam-se as exatidões globais das classificações a partir dos dados de treino extraídos da COS 2015 para as áreas de estudo A e B, respetivamente. Nas tabelas 12 e 14 apresentam-se as exatidões globais das generalizações das classificações anteriormente referidas para as áreas de estudo A e B, respetivamente. Para ambas as áreas, comparando os resultados na Tabela 11 com os resultados na Tabela 12 e os resultados na Tabela 13 com os resultados na Tabela 14, constata-se que a generalização usando um disco de raio 4 (pixéis) é benéfica de forma global, uma vez que a exatidão global aumenta com a generalização. Uma das possíveis explicações para o sucedido, é o facto de se estar a comparar o mapa criado com o mapa de referência cuja unidade mínima mapeável é de aproximadamente 1ha. Neste caso, o mapa de referência não terá uma UMC de 1ha, uma vez que este é criado a partir da conversão da COS 2018 em formato vetorial para matricial usando o critério Predominância (C1). Quando se realiza o processo de generalização reclassificam-se pixéis, ou seja, algumas áreas pequenas são reclassificadas como pertencentes a áreas vizinhas.

Observa-se que usando apenas as 4 bandas de 10m (F1) é menos vantajoso que usar um maior número de bandas. O uso adicional das bandas correspondentes aos índices radiométricas (F3) não mostraram vantagem em relação a usar apenas as bandas de S2 (F2). Por esta razão, dada a grande quantidade de classificações possíveis, nas análises subsequentes apresentar-se-á apenas os resultados usando o conjunto de bandas de S2 como *features*, F2. Pelo mesmo motivo, no uso de outras fontes de dados para se criarem dados de treino, subsecções seguintes, apenas se fará classificações com o conjunto de *features* F2.

Crítério\Features	F1	F2	F3
C0	84,7%	89,2%	88,7%
C1	84,9%	89,4%	88,8%
C2	84,7%	89,1%	88,6%

Tabela 11: Área de Estudo A – Exatidão global das classificações utilizando os dados de treino obtidos a partir da COS 2015.

Crítério\Features	F1	F2	F3
C0	87,1%	89,9%	89,4%
C1	87,2%	90,0%	89,4%
C2	87,1%	89,7%	89,3%

Tabela 12: Área de Estudo A – Exatidão global das generalizações das classificações utilizando os dados de treino obtidos a partir da COS 2015.

Critério\Features	F1	F2	F3
C0	70,4%	78,8%	77,4%
C1	70,7%	79,0%	77,6%
C2	70,4%	78,6%	77,2%

Tabela 13: Área de Estudo B – Exatidão global das classificações utilizando os dados de treino obtidos a partir da COS 2015.

Critério\Features	F1	F2	F3
C0	75,1%	80,9%	80,4%
C1	75,1%	80,6%	79,2%
C2	74,9%	80,3%	78,9%

Tabela 14: Área de Estudo B – Exatidão global das generalizações das classificações utilizando os dados de treino obtidos a partir da COS 2015.

Na Figura 24 e na Figura 25 são apresentados pormenores dos MUCSs obtidos para duas áreas mais pequenas contidas nas duas áreas de estudo. No Anexo F apresentam-se os mapas resultantes de todas as classificações e respetivas generalizações efetuadas com as amostras dos dados de treino para cada uma das áreas de estudo. Para cada MUCS são ainda apresentadas as correspondentes matrizes de confusão.

Como a classificação é feita ao nível do pixel, existem bastantes pixéis isolados, isto é, pixéis classificados como pertencentes a uma classe mas cujos vizinhos são classificados como pertencentes a outra(s) classe(s). Para a região representada na Figura 24, pertencente à área de estudo A, observa-se que a generalização efetuada diminuiu o número de pixéis isolados, mostrando ser um processo vantajoso de ser feito após a classificação, uma vez que o MUCS resultante é comparado com um mapa de referência criado a partir da COS 2018 com UMC de 1ha.

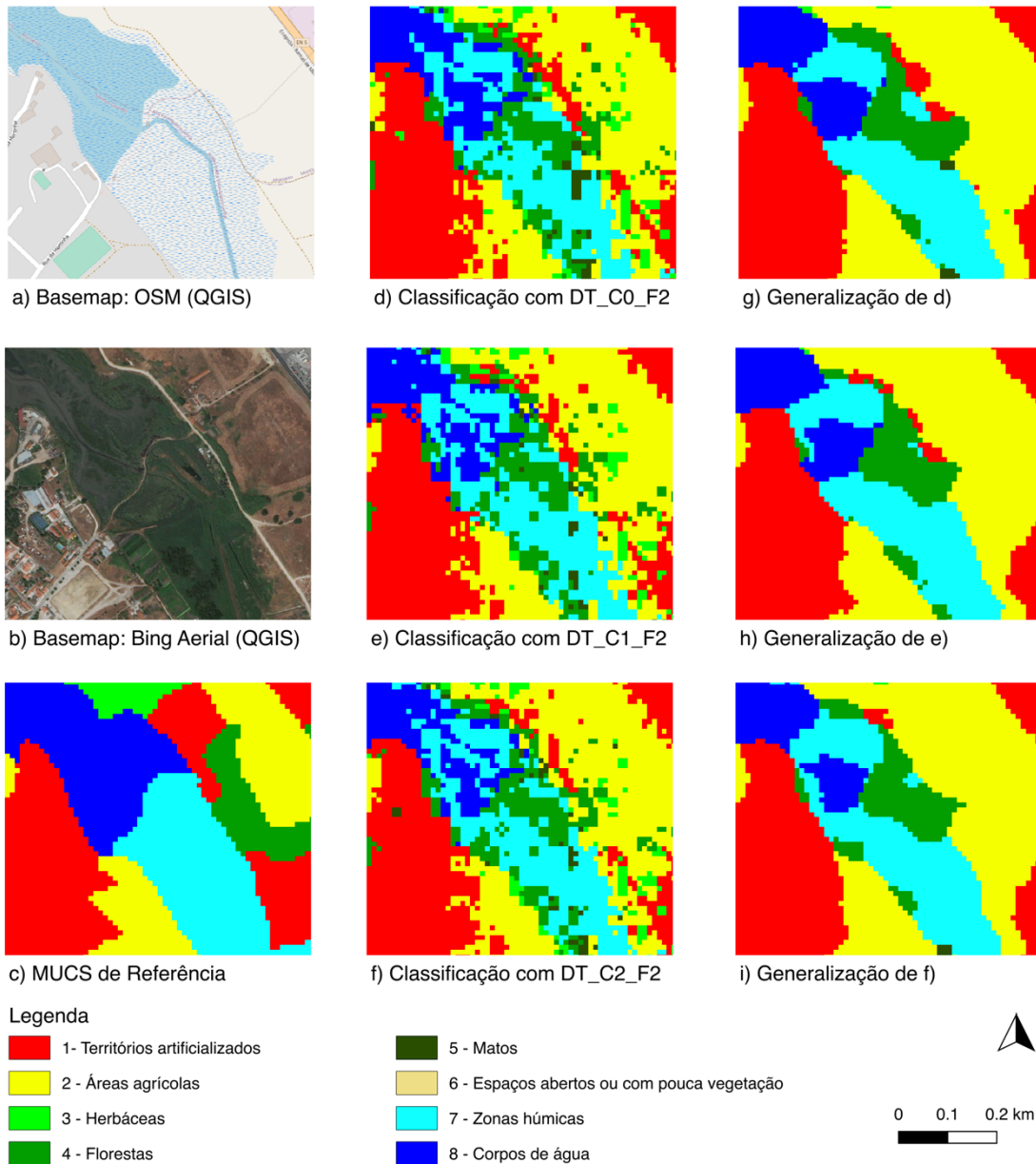


Figura 24: Região escolhida ao acaso na área de estudo A – MUCSs obtidos pelas classificações e respectivas generalizações.

No caso da região representada na Figura 25, pertencente à área de estudo B, observa-se o mesmo que na área de estudo A em relação à generalização, em comparação com o MUCS de referência. O detalhe da classificação ao nível pixel é perdido com a generalização. Nesta região observa-se uma grande confusão para a classificação nas classes 5 e 6.

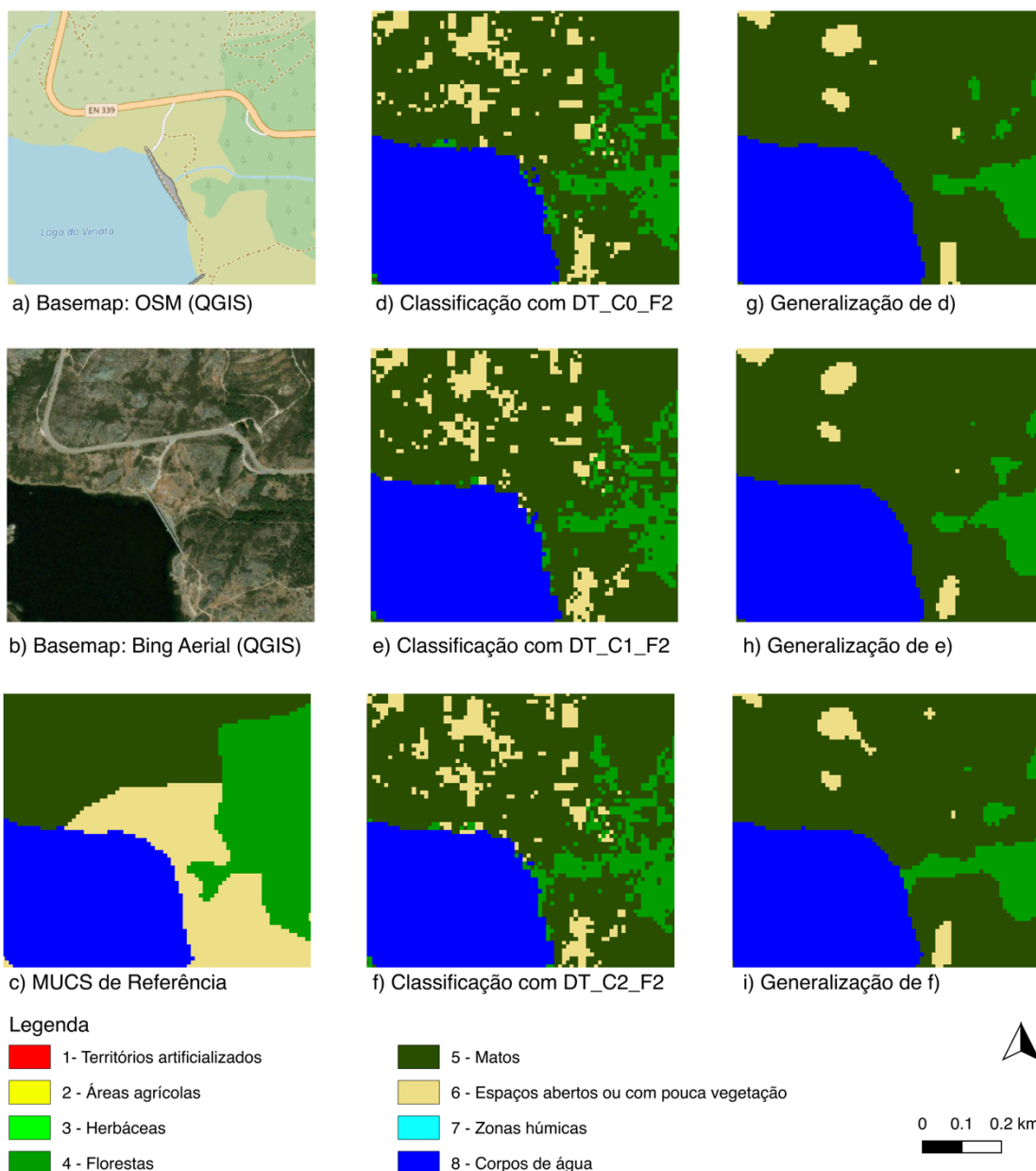


Figura 25: Região escolhida ao acaso na área de estudo B – MUCSs obtidos pelas classificações e respetivas generalizações.

Nas tabelas 16 e 17 são apresentadas, respetivamente, as exatidões do produtor e do utilizador dos MUCSs obtidos pela classificação e respetivas generalizações para a área de estudo A, e nas tabelas 18 e 19 são apresentados os dados correspondentes para a área de estudo B.

Para a área de estudo A, pode-se ver que na Tabela 15 a exatidão do utilizador diminui para as classes 1 e 3 com o processo de generalização, aumentando para as restantes classes. Quanto à exatidão do utilizador, para as classes 2, 3 e 6, o uso do critério C0 (Presença) mostrou ser o melhor critério de extração das áreas de treino a partir da COS 2015, sendo para as restantes melhor o critério C1 (Predominância).

Quanto à exatidão do produtor, pela Tabela 16, a generalização verifica ser sempre melhor para as classes 1, 2, 4, 7 e 8. Para todas as classes, o critério C1 (Predominância) apresenta melhores exatidões do produtor.

Independentemente do critério utilizado para criar as áreas de treino, destaca-se maiores erros de comissão para a classe 3 e de omissão para as classes 3, 5 e 6. Pelas matrizes de confusão nas tabelas 57, 58, 63, 64, 69 e 70 no Anexo F, os erros de omissão e comissão da classe 3 (Herbáceas) estão associados principalmente à classe 2 (Áreas agrícolas), dado que esta também contém vegetação rasteira que sofre modificações ao longo do ano. É de observar ainda que esta classe é a classe cujos dados de treino apresentavam a menor concordância com o mapa de referência feito a partir da COS 2018. A classe 5 (Matos) apresenta um maior erro de omissão associado às classes 1 (Territórios artificializados) e 4 (Superfícies agroflorestais e florestais), isto é, os pixels são classificados como pertencentes a 1 ou 4 mas no mapa de referência pertencem à classe 5, dado que a maioria das áreas de matos contém rochas e apresentam vegetação que se assemelha a árvores presentes na classe 4. Relativamente à classe 6 (Espaços descobertos ou com pouca vegetação), o seu erro de omissão está associado à classe 1 (Territórios artificializados) podendo estar associado ao facto que ambas as classes têm áreas contendo solo nu, ou seja, existem dados de treino para a classe 1 relativos a solo nu que fazem com que o classificador as classifique como pertencentes à classe 1 mas no mapa de referência pertencem à classe 6.

De uma forma geral, para a área de estudo A, o critério C1 (Presença) mostrou ser o que deu melhores resultados para a criação das áreas de treino a partir da COS 2015.

	Exatidão do Utilizador												
	DT_C0_F2			DT_C1_F2			DT_C2_F2			C1-C0		C2-C1	
	C	G	G - C	C	G	G - C	C	G	G - C	C	G	C	G
Classe 1	84,7%	83,3%	-1,4%	85,1%	83,5%	-1,6%	84,8%	83,3%	-1,4%	0,5%	0,3%	-0,4%	-0,2%
Classe 2	75,4%	78,3%	2,9%	75,3%	78,1%	2,8%	74,2%	77,0%	2,8%	-0,1%	-0,2%	-1,1%	-1,0%
Classe 3	55,4%	66,4%	11,0%	55,4%	66,6%	11,2%	54,9%	66,2%	11,3%	0,0%	0,1%	-0,5%	-0,3%
Classe 4	82,0%	83,1%	1,1%	82,2%	83,2%	1,0%	81,9%	83,0%	1,0%	0,2%	0,1%	-0,2%	-0,2%
Classe 5	77,8%	87,4%	9,6%	78,1%	87,6%	9,5%	77,2%	86,9%	9,8%	0,3%	0,2%	-1,0%	-0,7%
Classe 6	86,2%	92,9%	6,7%	85,8%	93,2%	7,3%	86,1%	93,0%	6,9%	-0,3%	0,2%	0,2%	-0,2%
Classe 7	91,7%	93,9%	2,2%	91,8%	94,0%	2,2%	91,4%	93,6%	2,2%	0,1%	0,1%	-0,4%	-0,4%
Classe 8	99,6%	99,5%	-0,1%	99,6%	99,5%	-0,1%	99,6%	99,5%	-0,1%	0,0%	0,0%	0,0%	0,0%

Tabela 15: Área de Estudo A: Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_C0_F2, DT_C1_F2 e DT_C2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

	Exatidão do Produtor												
	DT_C0_F2			DT_C1_F2			DT_C2_F2			C1-C0		C2-C1	
	C	G	G - C	C	G	G - C	C	G	G - C	C	G	C	G
Classe 1	92,6%	94,0%	1,4%	92,7%	94,0%	1,3%	92,4%	93,8%	1,3%	0,0%	0,0%	-0,3%	-0,3%
Classe 2	73,8%	77,1%	3,2%	74,3%	77,5%	3,2%	74,2%	77,4%	3,2%	0,4%	0,4%	-0,1%	0,0%
Classe 3	40,7%	37,2%	-3,5%	41,6%	37,6%	-4,0%	40,1%	36,5%	-3,5%	0,9%	0,4%	-1,6%	-1,1%
Classe 4	86,6%	88,0%	1,4%	86,8%	88,1%	1,4%	86,2%	87,7%	1,4%	0,2%	0,1%	-0,5%	-0,4%
Classe 5	44,6%	41,9%	-2,7%	45,2%	42,2%	-3,0%	44,1%	41,4%	-2,7%	0,6%	0,2%	-1,1%	-0,8%
Classe 6	55,1%	48,7%	-6,4%	56,2%	49,1%	-7,1%	53,8%	46,9%	-6,9%	1,1%	0,3%	-2,4%	-2,1%
Classe 7	79,7%	81,2%	1,6%	80,2%	81,7%	1,5%	79,2%	80,6%	1,3%	0,6%	0,5%	-1,0%	-1,1%
Classe 8	99,7%	99,8%	0,1%	99,7%	99,8%	0,1%	99,7%	99,7%	0,1%	0,0%	0,0%	0,0%	0,0%

Tabela 16: Área de Estudo A: Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_C0_F2, DT_C1_F2 e DT_C2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

Para a área de estudo B, pode-se ver que na Tabela 17 a exatidão do utilizador diminui para a classe 2 com o processo de generalização, aumentando para as restantes. Para as classes 3, e 8, o uso do critério C2 (Exclusividade) mostrou ser o melhor critério de extração das áreas de treino a partir da COS 2015, sendo para as restantes melhor o critério C1 (Predominância).

Quanto à exatidão do produtor, pela Tabela 18, a generalização verifica ser sempre melhor para as classes 2, 4, 5 e 6. Apenas para a classe 6, o uso do critério C0 (Presença) é o melhor, sendo para as restantes classes o melhor critério de extração das áreas de treino a partir da COS 2015 o critério C1 (Predominância).

Independentemente do critério utilizado para criar as áreas de treino, destaca-se maiores erros de comissão para a classe 3 e de omissão para as classes 3 e 8. Pelas matrizes de confusão nas tabelas 75, 76, 81, 82, 87 e 88 no Anexo F, tal e qual como na área de estudo A, os erros associados à classe 3 (Herbáceas) estão associados à classe 2 (Áreas agrícolas). O erro de omissão associado à classe 8 (Corpos de água) está associado a áreas que no mapa de referência pertencem à classe 4 (Superfícies agroflorestais e florestais), isto é, os pixéis são classificados como pertencentes à classe 4 mas no mapa de referência pertencem à classe 8.

De uma forma geral, para a área de estudo B, também o critério C1 (Presença) é aquele que deve ser considerado para a criação das áreas de treino a partir da COS 2015. Para ambas as áreas de estudo, destaca-se maiores erros de omissão e comissão para a classe 3 (Herbáceas).

	Exatidão do Utilizador												
	DT_C0_F2			DT_C1_F2			DT_C2_F2			C1-C0		C2-C1	
	C	G	G - C	C	G	G - C	C	G	G - C	C	G	C	G
Classe 1	80,7%	84,0%	3,3%	80,9%	83,7%	2,8%	80,6%	83,5%	2,9%	0,2%	-0,3%	-0,3%	-0,2%
Classe 2	78,4%	77,9%	-0,6%	78,6%	77,6%	-1,0%	78,2%	77,3%	-0,9%	0,2%	-0,2%	-0,5%	-0,4%
Classe 3	50,8%	73,0%	22,2%	50,5%	74,2%	23,7%	52,2%	77,3%	25,1%	-0,3%	1,2%	1,7%	3,1%
Classe 4	80,4%	82,1%	1,6%	80,6%	81,9%	1,3%	80,3%	81,6%	1,3%	0,2%	-0,1%	-0,3%	-0,3%
Classe 5	76,7%	80,6%	3,9%	76,9%	80,2%	3,3%	76,2%	79,7%	3,5%	0,2%	-0,4%	-0,7%	-0,5%
Classe 6	86,5%	91,4%	4,9%	86,8%	91,0%	4,2%	86,6%	91,1%	4,4%	0,2%	-0,4%	-0,1%	0,1%
Classe 8	91,6%	94,1%	2,5%	92,2%	93,5%	1,3%	92,7%	93,5%	0,8%	0,7%	-0,6%	0,4%	-0,1%

Tabela 17: Área de Estudo B: Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_C0_F2, DT_C1_F2 e DT_C2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

	Exatidão do Produtor												
	DT_C0_F2			DT_C1_F2			DT_C2_F2			C1-C0		C2-C1	
	C	G	G - C	C	G	G - C	C	G	G - C	C	G	C	G
Classe 1	63,6%	62,8%	-0,8%	65,0%	62,6%	-2,4%	62,8%	61,0%	-1,8%	1,4%	-0,2%	-2,3%	-1,6%
Classe 2	84,1%	86,8%	2,7%	84,4%	86,8%	2,4%	84,1%	86,4%	2,3%	0,3%	0,0%	-0,3%	-0,4%
Classe 3	20,9%	12,5%	-8,5%	21,6%	11,8%	-9,8%	20,1%	11,3%	-8,8%	0,6%	-0,7%	-1,5%	-0,5%
Classe 4	84,1%	86,2%	2,1%	84,3%	85,9%	1,6%	83,9%	85,6%	1,7%	0,2%	-0,3%	-0,4%	-0,3%
Classe 5	76,0%	78,9%	2,9%	75,9%	78,7%	2,8%	75,8%	78,5%	2,7%	-0,1%	-0,3%	-0,1%	-0,2%
Classe 6	65,9%	67,6%	1,7%	65,9%	67,3%	1,4%	65,4%	67,1%	1,7%	0,0%	-0,2%	-0,5%	-0,2%
Classe 8	57,6%	52,2%	-5,4%	59,0%	51,8%	-7,2%	56,2%	50,6%	-5,6%	1,4%	-0,4%	-2,8%	-1,2%

Tabela 18: Área de Estudo B: Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_C0_F2, DT_C1_F2 e DT_C2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

Na Figura 26 e na Figura 27 são apresentadas as áreas cujo uso e ocupação do solo sofreram alterações da COS 2015 para a COS 2018. Para a área de estudo A, tem-se que 3,4% da área total sofreu mudança. No caso da área de estudo B, esta tem 6,9% da sua área mudada na atualização da COS 2015 para a COS 2018.

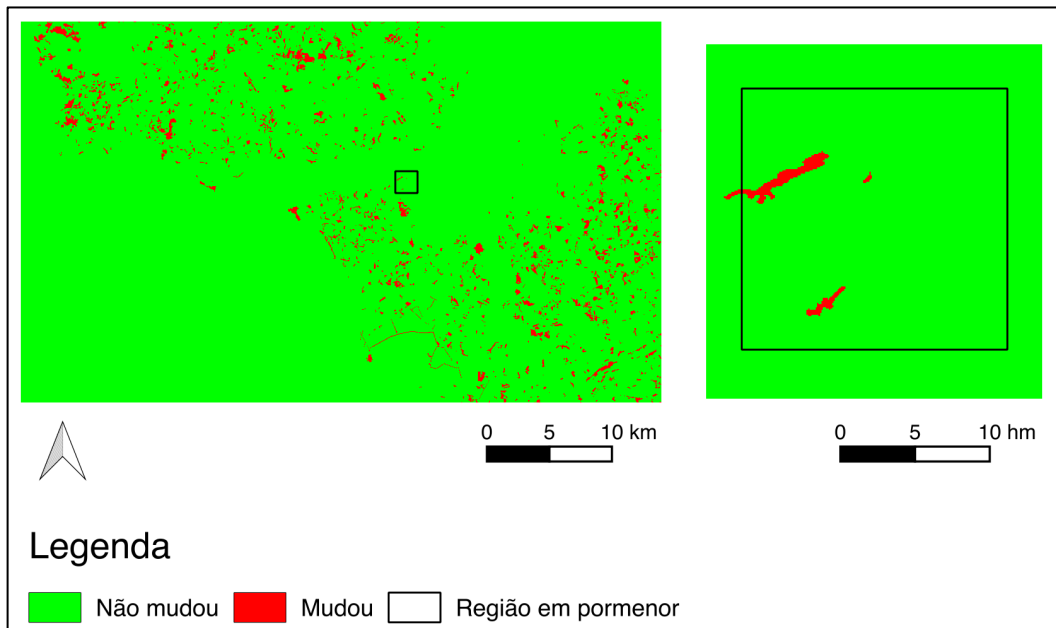


Figura 26: Área de estudo A: Mudança de uso e ocupação do solo entre a COS 2015 e a COS 2018.

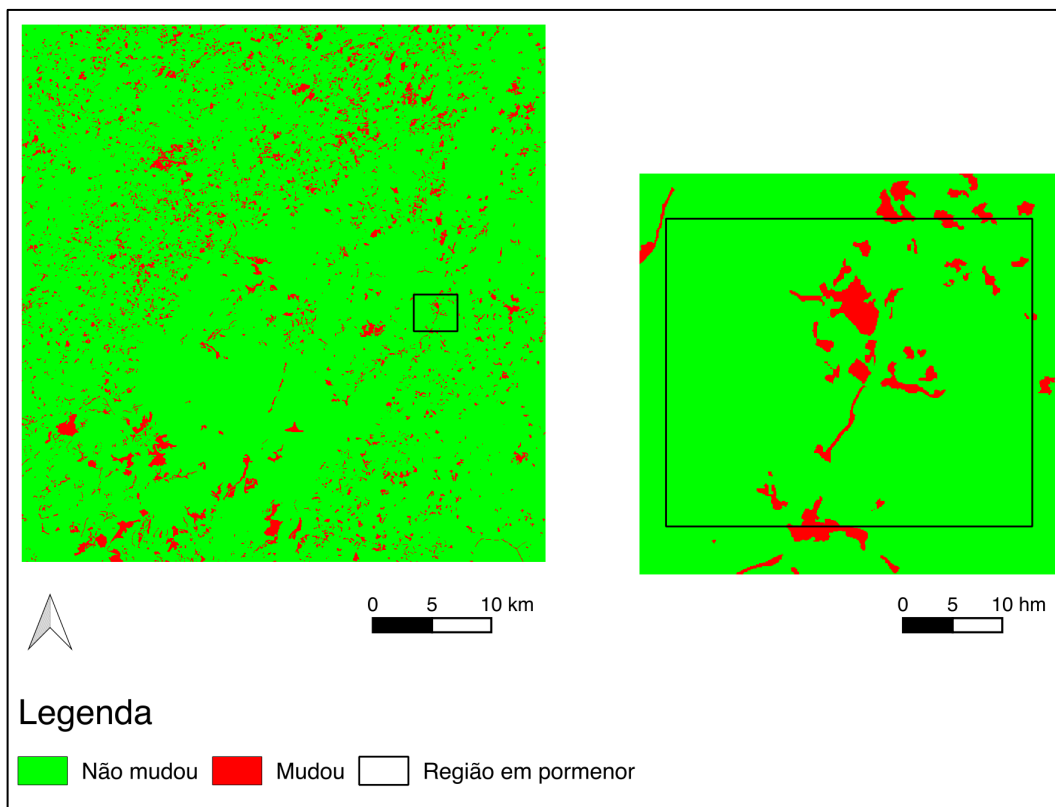


Figura 27: Área de estudo B: Mudança de uso e ocupação do solo entre a COS 2015 e a COS 2018.

Para as áreas de estudo A e B, são apresentadas nas tabelas 20 e 21, respectivamente, as exatidões globais avaliadas apenas para as áreas onde ocorreram mudanças de uso e ocupação do solo entre a COS 2015 e a COS 2018. Verifica-se que com as generalizações efetuadas nos MUCSs obtidos pelas classificações apresentam melhores resultados nas áreas onde ocorreram mudança de ocupação e uso do solo. Para ambas as áreas de estudo, verifica-se que se obteve melhores resultados utilizando o conjunto de *features* F1 para as áreas de mudança.

	Área de Mudança					
	Classificação			Generalização		
	F1	F2	F3	F1	F2	F3
C0	37,5%	34,2%	36,6%	42,3%	37,4%	39,9%
C1	37,3%	34,0%	36,4%	42,6%	37,7%	40,2%
C2	37,8%	35,0%	37,3%	42,5%	38,2%	40,6%

Tabela 19: Área de estudo A: Percentagem de área corretamente classificada nas áreas onde ocorreram mudança de ocupação e uso do solo, da COS 2015 para COS 2018.

	Área de Mudança					
	Classificação			Generalização		
	F1	F2	F3	F1	F2	F3
C0	44,4%	38,9%	41,5%	48,8%	42,6%	44,8%
C1	44,0%	38,7%	41,3%	48,9%	42,9%	44,9%
C2	44,7%	39,6%	42,0%	49,0%	43,2%	45,1%

Tabela 20: Área de estudo B: Percentagem de área corretamente classificada nas áreas onde ocorreram mudança de ocupação e uso do solo, da COS 2015 para COS 2018.

4.2.2. Classificação com dados de treino extraídos da COS 2015 com filtragens a partir dos índices radiométricos

Como o critério C1 (Presença) para criação das áreas de treino a partir da COS 2015 mostrou ser o melhor, utiliza-se estas áreas para se testarem filtragens com os índices radiométricos. Os novos dados de treino são assim obtidos a partir destas áreas em conjunto com as *features* F2, uma vez que foi com este conjunto de *features* que se obteve MUCSs mais próximos do mapa de referência criado a partir da COS 2018.

Para utilizar o critério L, selecionou-se os índices NDVI, NDWI e NDBI, e a partir da literatura e da análise visual e dos histogramas por classe (disponíveis no Anexo G) definiram-se as condições reunidas na Tabela 21.

Para a classe 1, territórios artificializados, a ausência de água em qualquer momento do ano permite inferir que o NDWI seja negativo para as três imagens. A presença de vegetação nesta classe deverá ser muito reduzida, ou mesmo não existir, tendo-se assim estabelecido o NDVI inferior a 0.3 em qualquer altura do ano. Como esta classe deverá ser predominantemente construções, determinou-se que o NDBI deverá ser positivo em pelo menos uma das datas.

Para as classes de vegetação, classes 2, 3, 4 e 5, dada a presença de vegetação e ausência de água, obriga-se que o índice NDVI seja superior a 0.3 e o índice NDWI seja negativo para qualquer uma das imagens.

As classes 6 e 7, referentes a espaços descobertos ou com pouca vegetação e zonas húmidas, respetivamente, são classes cujo comportamento ao longo de um ano varia bastante, impõem-se condições para pelo menos uma das imagens. No caso da classe 6, dado que se espera que em pelo menos uma das alturas do ano haja vegetação e não haja água, impõem-se que o índice NDVI seja positivo e o índice NDWI seja negativo. Com a classe 7, a presença de vegetação deverá ser escassa e com a possibilidade de ausência de água à superfície, obriga-se as mesmas condições impostas para a classe 6.

Com a classe 8, referente aos corpos de água, a presença de água ao longo do ano é obrigatória, impondo-se assim que o índice NDWI seja positivo. Como na água poderá haver presença de algas, não esperando que seja em grande quantidade, obriga-se que o índice NDVI seja inferior a 0.3 em pelo menos uma das imagens.

Nomenclatura a adotar	NDVI / imagens	NDWI / imagens	NDBI / imagens
1 Territórios artificializados	<0.3 / todas	<0.0 / todas	>0.0 / pelo menos uma
2 Áreas agrícolas	>0.3 / todas	<0.0 / todas	---
3 Herbáceas	>0.3 / todas	<0.0 / todas	---
4 Superfícies agroflorestais e florestais	>0.3 / todas	<0.0 / todas	---
5 Matos	>0.3 / todas	<0.0 / todas	---
6 Espaços descobertos ou com pouca vegetação	>0.0 / pelo menos uma	<0.0 / pelo menos uma	---
7 Zonas húmidas	>0.0 / pelo menos uma	<0.0 / pelo menos uma	---
8 Corpos de água	<0.3 / pelo menos uma	>0.0 / todas	---

Tabela 21: Condições a verificar por cada classe a partir dos índices radiométricos NDVI, NDWI e NDBI.

Na Tabela 22 e na Tabela 23 são apresentadas as comparações das amostras dos dados de treino criados para as áreas de estudo A e B, respectivamente, com o mapa de referência. Verifica-se que para a área de estudo A, de forma global, a qualidade dos dados de treino aumenta com ambas as filtragens. Para as classes 3 e 6, a percentagem de pixéis corretos diminui quer usando o critério L quer usando o critério M. No caso do uso do critério L, também a qualidade diminui para as classes 2 e 6. Quanto à área de estudo B, de forma global apenas o critério M mostrou ser vantajoso, diminuindo apenas a qualidade dos dados de treino para a classe 3. Com o critério L, tanto a qualidade global dos dados de treino, como em particular para as classes 2, 3, e 6, as qualidades diminuiram.

	S_DT_C1		S_DT_C1L		S_DT_C1M	
	Nº Pixéis	% Corretos	Nº Pixéis	% Corretos	Nº Pixéis	% Corretos
Classe 1	373943	98,7%	175712	99,8%	369045	98,7%
Classe 2	132170	95,4%	45186	93,1%	130821	95,4%
Classe 3	68389	63,2%	34797	62,7%	67391	63,0%
Classe 4	213071	95,9%	177175	97,2%	210702	96,2%
Classe 5	63845	88,9%	43007	90,6%	62446	89,3%
Classe 6	6859	87,5%	6738	86,4%	6489	86,2%
Classe 7	12372	99,8%	12213	99,8%	12000	99,8%
Classe 8	688895	100,0%	656855	100,0%	668704	100,0%
Global	1559544	96,6%	1151683	97,7%	1527598	96,6%

Tabela 22: Área de estudo A – Validação das amostras dos dados de treino criados a partir da filtragem com os índices radiométricos das áreas de treino DT_C1: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos. Denota-se por S_DT_C1, S_DT_C1L e S_DT_C1M as amostras de DT_C1, DT_C1L e DT_C1M, respetivamente.

	S_DT_C1		S_DT_C1L		S_DT_C1M	
	Nº Pixéis	% Corretos	Nº Pixéis	% Corretos	Nº Pixéis	% Corretos
Classe 1	62670	97,4%	20161	99,0%	62321	97,6%
Classe 2	431240	95,2%	310583	94,7%	428153	95,2%
Classe 3	89471	42,0%	60693	34,6%	85936	40,5%
Classe 4	773047	95,9%	601646	96,6%	761935	95,9%
Classe 5	558805	94,4%	335212	93,5%	557319	94,4%
Classe 6	78595	98,5%	78590	98,5%	77790	98,5%
Classe 7	0	-	0	-	0	-
Classe 8	5708	99,4%	1224	100,0%	5528	99,4%
Global	1999536	93,1%	1408109	92,9%	1978982	93,1%

Tabela 23: Área de estudo B – Validação das amostras dos dados de treino criados a partir da filtragem com os índices radiométricos das áreas de treino DT_C1: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos. Denota-se por S_DT_C1, S_DT_C1L e S_DT_C1M as amostras de DT_C1, DT_C1L e DT_C1M, respetivamente.

Nos gráficos da Figura 28 e da Figura 29, apresenta-se a separabilidade dos dados de treino criados a partir da filtragem de DT_C1_F2 utilizando os critérios L e M. A filtragem com o critério L mostrou aumentar a separabilidade para todas as classes, exceto para a classe 7 na área de estudo A e para a classe 2 na área de estudo B. No caso do uso do critério M, a separabilidade aumentou para todas as classes em ambas as áreas de estudo.

Utilizando a informação da separabilidade e da comparação dos dados de treino com o mapa de referência, não existe nenhuma relação direta ou indireta do aumento de separabilidade com o aumento ou diminuição da qualidade dos dados de treino.

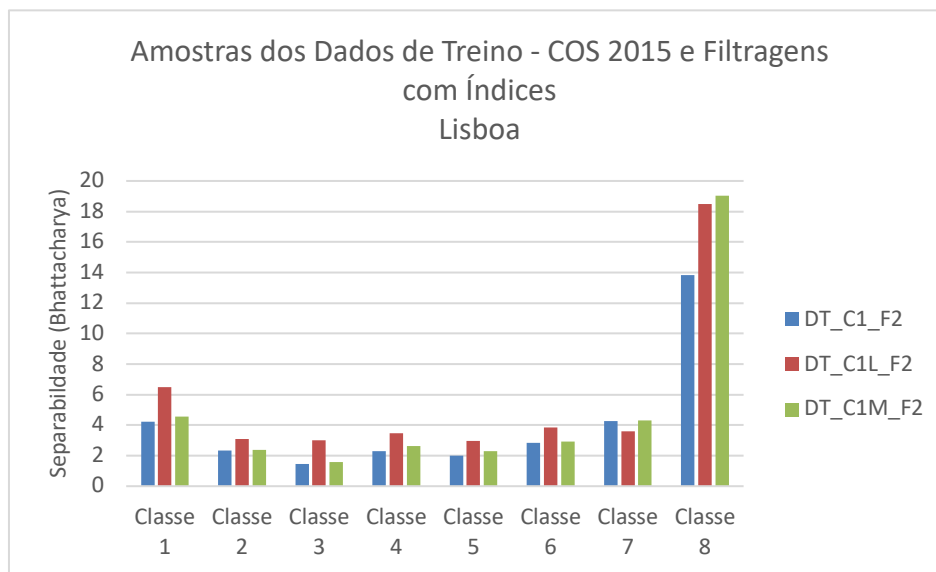


Figura 28: Área de estudo A – Avaliação da separabilidade da amostra de dados de treino obtidos através da COS 2015 com filtragens a partir dos índices radiométricos usando a distância de Bhattacharya.

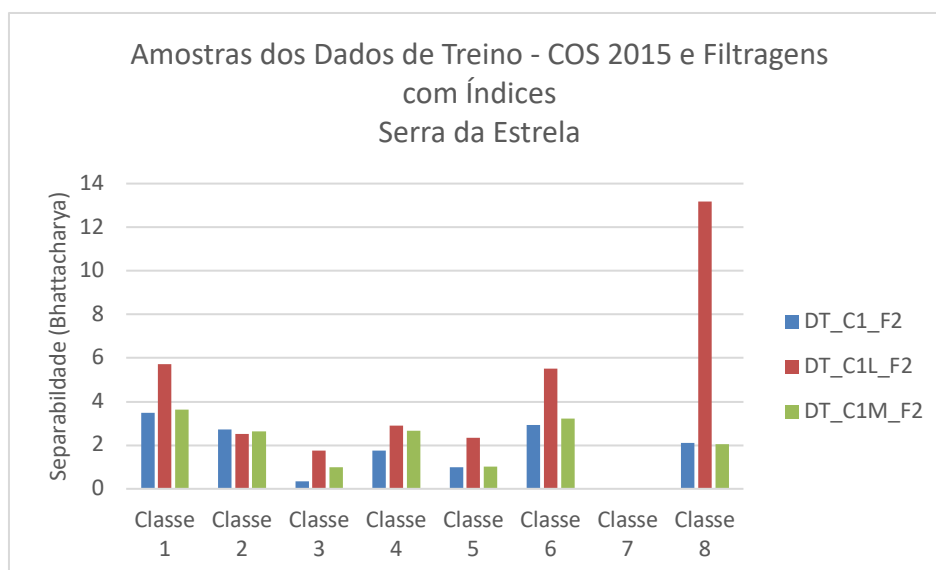


Figura 29: Área de estudo B – Avaliação da separabilidade da amostra de dados de treino obtidos através da COS 2015 com filtragens a partir dos índices radiométricos usando a distância de Bhattacharya.

Na Figura 30 e na Figura 31 são apresentadas os MUCSs obtidos para duas áreas mais pequenas contidas nas duas áreas de estudo usando os dados de treino DT_C1_F2, DT_C1L_F2 e DT_C1M_F2. No Anexo H apresentam-se todas as classificações e respetivas generalizações efetuadas com as amostras dos dados de treino para ambas as áreas de estudo. São também apresentadas as matrizes de confusão de cada MUCS obtido.

Na Figura 30, para a região escolhida ao acaso da área de estudo A, observa-se de imediato que os corpos de água deixaram de ser classificados com o uso dos dados de treino filtrados a partir dos índices radiométricos. Com estes dados de treino, observa-se principalmente que passou a existir uma maior confusão na classificação entre as classes 1 e 2 e entre as classes 2 e 3.

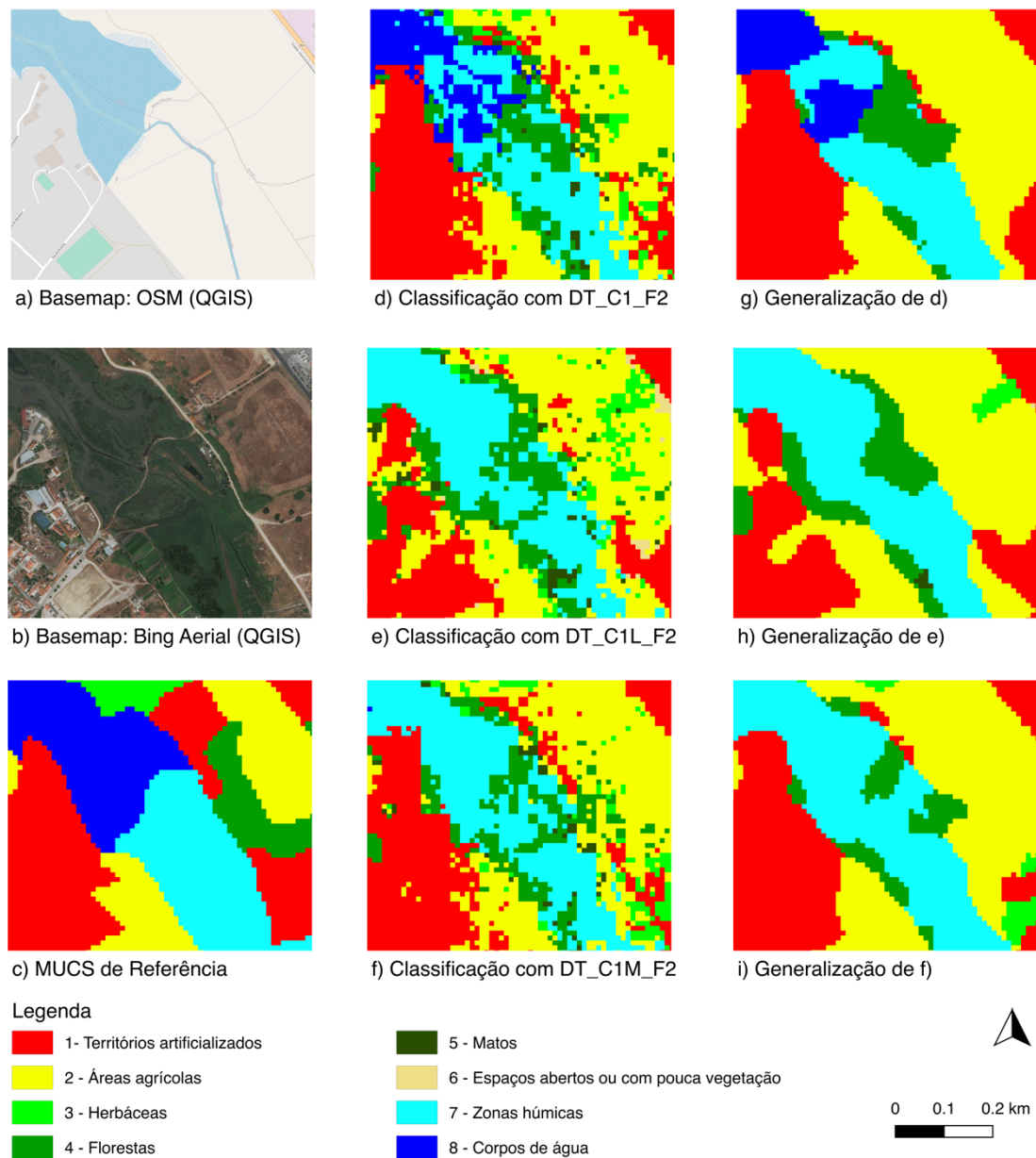


Figura 30: Região escolhida ao acaso na área de estudo A – MUCSs obtidos pelas classificações e respetivas generalizações a partir dos dados de treino criados com filtragens usando os índices radiométricos.

Na Figura 31, para a região escolhida ao acaso da área de estudo B, observa-se de imediato uma grande diferença entre a classificação feita com DT_C1_F2 e DT_C1L_F2, passando a classe 5 a estar quase inexistente e havendo confusão na classificação com a classe 4. Com os dados de treino DT_C1M_F2, observa-se uma maior confusão entre as classes 4 e 5 que não era tão grande com o uso dos dados de treino DT_C1_F2.

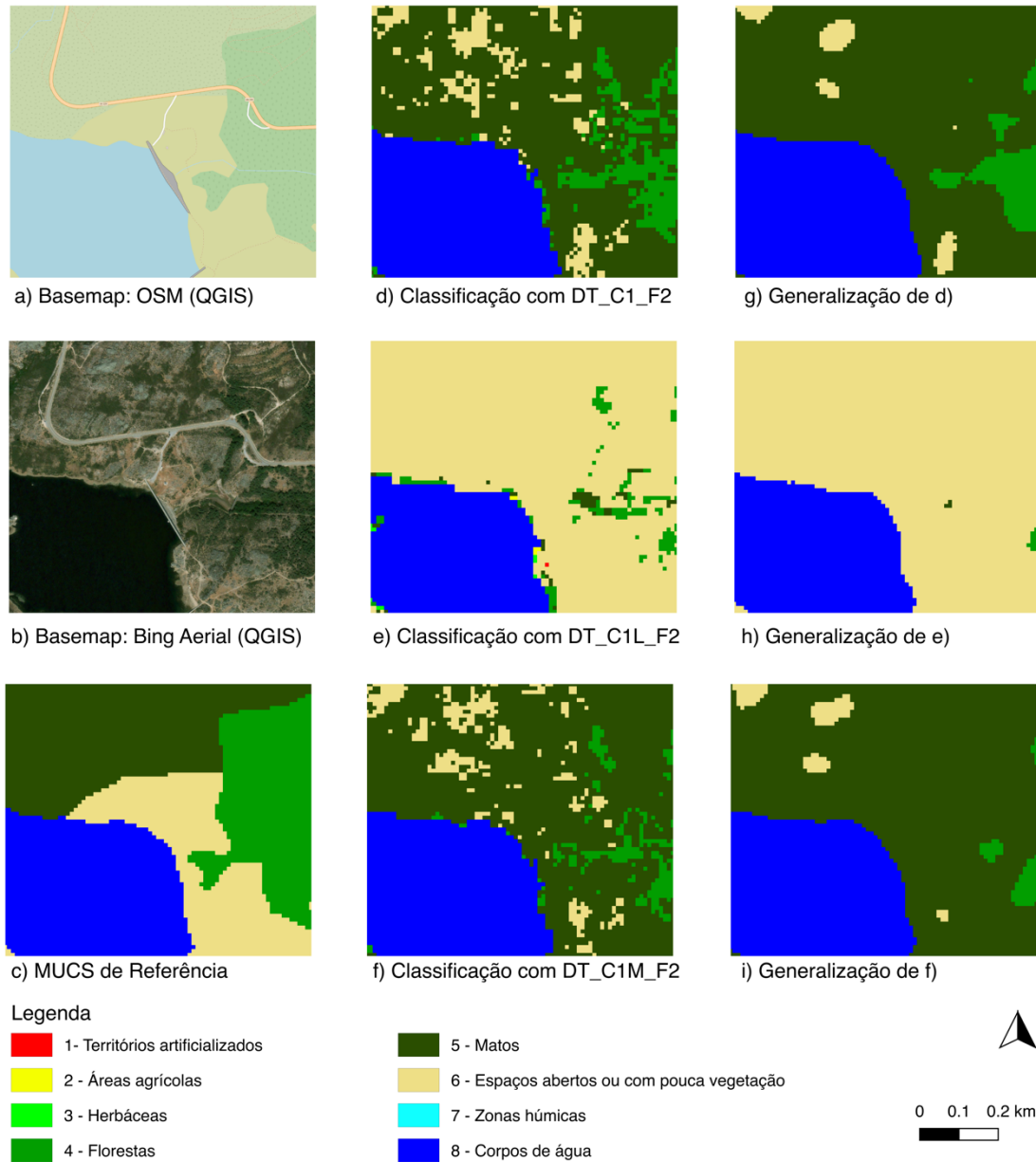


Figura 31: Região escolhida ao acaso na área de estudo B – MUCSs obtidos pelas classificações e respectivas generalizações a partir dos dados de treino criados com filtragens usando os índices radiométricos.

Na Tabela 24 têm-se as exatidões globais das classificações obtidas com os dados de treino extraídos a partir da COS 2015 com filtragem utilizando os índices radiométricos com os critérios L e M. Pode-se observar que entre os dois critérios, o uso do critério M como forma de filtragem dos dados de treino provenientes da COS 2015 com o critério C1 (Predominância), apresentou melhores resultados de classificação em relação ao critério L, mas não superou o facto de não ser feita nenhuma filtragem, ou seja, utilizando os dados de treino DT_C1_F2.

Os MUCSs criados a partir dos dados de treino obtidos por filtragem de DT_C1_F2 com recurso aos índices radiométricos, verifica-se que qualquer generalização feita apresenta sempre melhorias em relação à classificação.

	Área de Estudo A		Área de Estudo B	
	Classificação	Generalização	Classificação	Generalização
DT_C1_F2	89,4%	90,0%	79,0%	80,6%
DT_C1L_F2	81,5%	85,0%	69,3%	71,6%
DT_C1M_F2	88,6%	89,1%	78,7%	80,3%

Tabela 24: Exatidões globais das classificações, e respetivas generalizações, utilizando os dados de treino DT_C1_F2 com filtragens utilizando os índices radiométricos com os critérios L e M.

Nas duas tabelas seguintes, Tabela 25 e Tabela 26, referentes à área de estudo A, são apresentadas as exatidões do utilizador e do produtor por classe, respetivamente. Para a exatidão do utilizador apenas existe melhoria para a classe 8 e no caso da exatidão do produtor para as classes 6 e 7. Quanto aos valores de cada exatidão para cada classe mantém-se semelhantes, destacando-se maiores erros de comissão para a classe 3 e de omissão para as classes 3, 5 e 6.

	Exatidão do Utilizador													
	DT_C1_F2			DT_C1L_F2			DT_C1M_F2			C1L-C0		C1M-C1		
	C	G	G - C	C	G	G - C	C	G	G - C	C	G	C	G	
Classe 1	84,7%	83,3%	-1,4%	85,1%	83,5%	-1,6%	84,8%	83,3%	-1,4%	0,5%	0,3%	-0,4%	-0,2%	
Classe 2	75,4%	78,3%	2,9%	75,3%	78,1%	2,8%	74,2%	77,0%	2,8%	-0,1%	-0,2%	-1,1%	-1,0%	
Classe 3	55,4%	66,4%	11,0%	55,4%	66,6%	11,2%	54,9%	66,2%	11,3%	0,0%	0,1%	-0,5%	-0,3%	
Classe 4	82,0%	83,1%	1,1%	82,2%	83,2%	1,0%	81,9%	83,0%	1,0%	0,2%	0,1%	-0,2%	-0,2%	
Classe 5	77,8%	87,4%	9,6%	78,1%	87,6%	9,5%	77,2%	86,9%	9,8%	0,3%	0,2%	-1,0%	-0,7%	
Classe 6	86,2%	92,9%	6,7%	85,8%	93,2%	7,3%	86,1%	93,0%	6,9%	-0,3%	0,2%	0,2%	-0,2%	
Classe 7	91,7%	93,9%	2,2%	91,8%	94,0%	2,2%	91,4%	93,6%	2,2%	0,1%	0,1%	-0,4%	-0,4%	
Classe 8	99,6%	99,5%	-0,1%	99,6%	99,5%	-0,1%	99,6%	99,5%	-0,1%	0,0%	0,0%	0,0%	0,0%	

Tabela 25: Área de Estudo A – Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_C1_F2, DT_C1L_F2 e DT_C1M_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

Exatidão do Produtor													
	DT_C1_F2			DT_C1L_F2			DT_C1M_F2			C1L-C1		C1M-C1	
	C	G	G - C	C	G	G - C	C	G	G - C	C	G	C	G
Classe 1	92,6%	94,0%	1,4%	92,7%	94,0%	1,3%	92,4%	93,8%	1,3%	0,0%	0,0%	-0,3%	-0,3%
Classe 2	73,8%	77,1%	3,2%	74,3%	77,5%	3,2%	74,2%	77,4%	3,2%	0,4%	0,4%	-0,1%	0,0%
Classe 3	40,7%	37,2%	-3,5%	41,6%	37,6%	-4,0%	40,1%	36,5%	-3,5%	0,9%	0,4%	-1,6%	-1,1%
Classe 4	86,6%	88,0%	1,4%	86,8%	88,1%	1,4%	86,2%	87,7%	1,4%	0,2%	0,1%	-0,5%	-0,4%
Classe 5	44,6%	41,9%	-2,7%	45,2%	42,2%	-3,0%	44,1%	41,4%	-2,7%	0,6%	0,2%	-1,1%	-0,8%
Classe 6	55,1%	48,7%	-6,4%	56,2%	49,1%	-7,1%	53,8%	46,9%	-6,9%	1,1%	0,3%	-2,4%	-2,1%
Classe 7	79,7%	81,2%	1,6%	80,2%	81,7%	1,5%	79,2%	80,6%	1,3%	0,6%	0,5%	-1,0%	-1,1%
Classe 8	99,7%	99,8%	0,1%	99,7%	99,8%	0,1%	99,7%	99,7%	0,1%	0,0%	0,0%	0,0%	0,0%

Tabela 26: Área de Estudo A – Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_C1_F2, DT_C1L_F2 e DT_C1M_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

Na Tabela 27 e na Tabela 28, referentes à área de estudo B, são apresentadas as exatidões do utilizador e do produtor por classe, respetivamente. Para esta área de estudo, verifica-se uma melhoria significativa na exatidão do produtor para a classe 6 usando os dados de treino criados a partir da filtragem usando o critério L, DT_C1L_F2. Usando estes mesmos dados de treino, a classe 8 apresenta uma melhoria quanto à exatidão do utilizador. Com o critério M, a exatidão do utilizador aumenta muito pouco para a classes 2 e 4, e a exatidão do produtor aumenta para as classes 1, 5 e 8. Com qualquer um dos critérios, as classes destacadas com maiores erros de omissão e comissão mantêm-se as mesmas que não usando qualquer tipo de filtragem a partir dos índices radiométricos.

Exatidão do Utilizador													
	DT_C1_F2			DT_C1L_F2			DT_C1M_F2			C1L-C1		C1M-C1	
	C	G	G - C	C	G	G - C	C	G	G - C	C	G	C	G
Classe 1	80,9%	83,7%	2,8%	43,1%	51,1%	8,0%	78,8%	82,2%	3,3%	-37,8%	-32,7%	-2,1%	-1,6%
Classe 2	78,6%	77,6%	-1,0%	76,6%	75,7%	-0,8%	78,8%	77,7%	-1,1%	-2,1%	-1,9%	0,1%	0,1%
Classe 3	50,5%	74,2%	23,7%	41,4%	68,3%	26,9%	47,6%	70,0%	22,4%	-9,1%	-5,9%	-2,9%	-4,2%
Classe 4	80,6%	81,9%	1,3%	79,5%	80,7%	1,2%	80,6%	81,9%	1,2%	-1,1%	-1,2%	0,0%	0,0%
Classe 5	76,9%	80,2%	3,3%	72,4%	77,2%	4,8%	75,9%	79,1%	3,1%	-4,5%	-3,0%	-1,0%	-1,1%
Classe 6	86,8%	91,0%	4,2%	27,9%	29,0%	1,1%	86,4%	91,1%	4,8%	-58,8%	-62,0%	-0,4%	0,2%
Classe 8	92,2%	93,5%	1,3%	95,4%	96,2%	0,8%	86,2%	88,3%	2,1%	3,1%	2,6%	-6,0%	-5,2%

Tabela 27: Área de Estudo B – Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_C1_F2, DT_C1L_F2 e DT_C1M_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

	Exatidão do Produtor												
	DT_C1_F2			DT_C1L_F2			DT_C1M_F2			C1L-C1		C1M-C1	
	C	G	G - C	C	G	G - C	C	G	G - C	C	G	C	G
Classe 1	65,0%	62,6%	-2,4%	52,7%	52,1%	-0,7%	66,0%	63,7%	-2,2%	-12,3%	-10,5%	1,0%	1,2%
Classe 2	84,4%	86,8%	2,4%	79,1%	82,7%	3,6%	84,2%	86,6%	2,4%	-5,3%	-4,1%	-0,2%	-0,1%
Classe 3	21,6%	11,8%	-9,8%	11,1%	4,5%	-6,6%	18,4%	8,8%	-9,6%	-10,4%	-7,3%	-3,1%	-3,0%
Classe 4	84,3%	85,9%	1,6%	78,4%	80,7%	2,3%	83,6%	85,1%	1,5%	-5,9%	-5,2%	-0,7%	-0,8%
Classe 5	75,9%	78,7%	2,8%	54,0%	56,4%	2,4%	76,2%	78,7%	2,6%	-21,9%	-22,2%	0,3%	0,1%
Classe 6	65,9%	67,3%	1,4%	85,4%	86,5%	1,1%	65,5%	66,9%	1,4%	19,5%	19,2%	-0,4%	-0,4%
Classe 8	59,0%	51,8%	-7,2%	37,2%	37,8%	0,7%	61,0%	53,0%	-8,0%	-21,9%	-14,0%	2,0%	1,2%

Tabela 28: Área de Estudo B – Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_C1_F2, DT_C1L_F2 e DT_C1M_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

Tem-se assim em comum para ambas as áreas de estudo, que a exatidão do utilizador aumenta com o a filtragem usando o critério L e a exatidão do produtor aumenta para a classe 6 usando o mesmo critério de filtragem.

4.2.3. Classificação com dados de treino extraídos do OSM

Dado que as exatidões globais usando todo o conjunto de bandas de S2, F2, apresentou melhores resultados para ambas as áreas de estudo, para o conjunto de dados provenientes do OSM utiliza-se o conjunto de *features* F2 para a criação de dados de treino.

O aumento da separabilidade das classes, assim como da sua concordância com a COS 2018, tanto utilizando os dados diretamente provenientes da COS 2015, como com filtragens utilizando os índices radiométricos, não mostraram ter uma relação direta ou indireta com as exatidões globais, do utilizador e do produtor. Desta forma, nesta secção e seguintes, não se apresenta a avaliação dos dados de treino a partir da separabilidade.

Como existem áreas sem classificação por parte do OSM, utilizando os critérios O0 (Presença), O1 (Predominância) e O2 (Exclusividade) na criação dos dados de treino provenientes deste produto, poderão existir áreas interiores nas áreas de estudo que não serão cobertas por nenhuns dados de treino.

Na Figura 32 e na Figura 33 são apresentadas as áreas cobertas pelos dados de treino criados a partir do OSM para a área de estudo A e B, respetivamente. A amarelo tem-se as áreas cobertas por DT_O2, e a azul a área coberta por DT_O1 e DT_O0 exceto DT_O2. Assim, a área coberta por

DT_O1 coincide com DT_O0 que contém DT_O2. Para a área de estudo A, referente a Lisboa, excluindo região do oceano, a quantidade de cobertura do OSM é muito maior que a da área de estudo B, referente ao Parque Natural da Serra da Estrela.

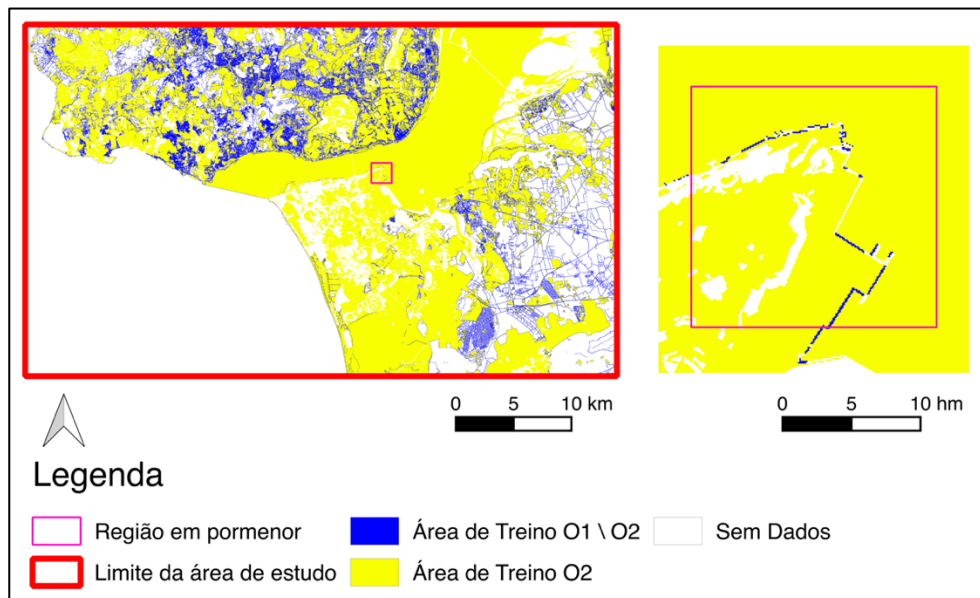


Figura 32: Área de estudo A – Áreas cobertas pelos dados de treino obtidos a partir do OSM (azul – área coberta por DT_O0 e DT_O1; amarelo – área pertencente a DT_O0, DT_O1 e DT_O2).

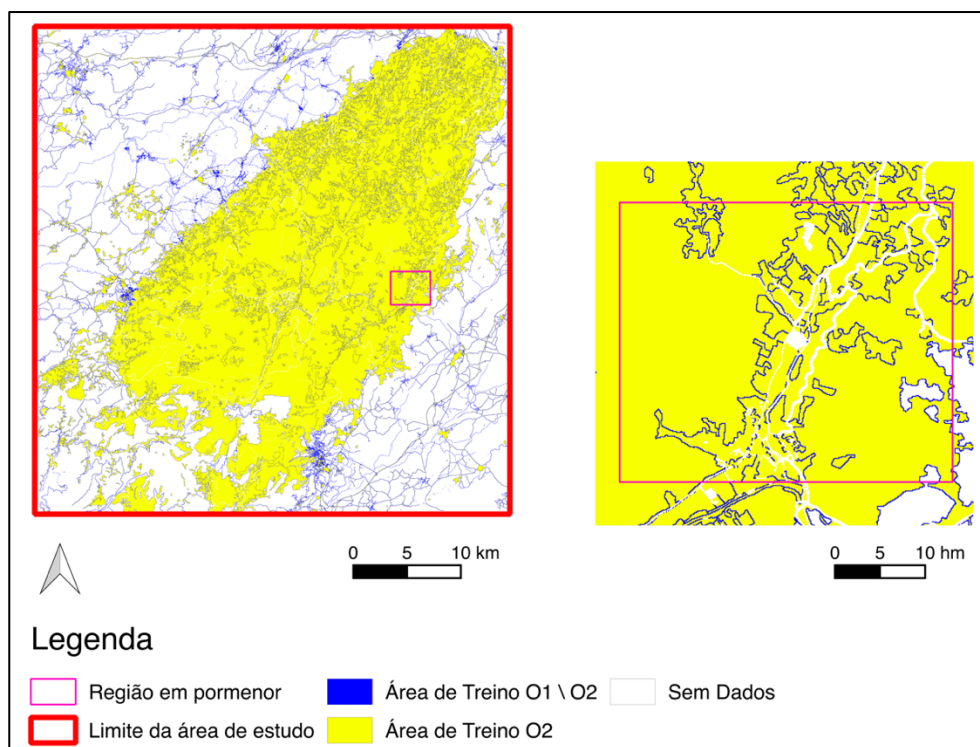


Figura 33: Área de estudo B – Áreas cobertas pelos dados de treino obtidos a partir do OSM (azul – área coberta por DT_O0 e DT_O1; amarelo – área pertencente a DT_O0, DT_O1 e DT_O2).

Nas tabelas 29 e 30 são apresentadas as comparações dos dados de treino criados para a área de estudo A e para a área de estudo B, respectivamente. Em geral, comparando a quantidade de dados de treino provenientes do OSM com a COS 2015, para cada classe tem-se mais dados de treino com a COS 2015, exceto para a classe 1 na área de estudo B. Um dos motivos está relacionado com a generalização que é feita na produção da COS 2015 de forma a UMC ser 1ha. Assim os territórios urbanos com ocupação menor que 1ha são generalizados para outras classes. Como o processo de produção da COS 2018 é semelhante, a percentagem de dados classificados corretamente também é superior com a COS 2015. De forma geral, os melhores dados de treino obtidos com o OSM são criados utilizando o critério O2 (Exclusividade).

	S_DT_O0		S_DT_O1		S_DT_O2	
	Nº Pixéis	% Corretos	Nº Pixéis	% Corretos	Nº Pixéis	% Corretos
Classe 1	356405	80,8%	353172	81,0%	253122	81,3%
Classe 2	19980	67,9%	19297	69,3%	15396	72,4%
Classe 3	72445	10,3%	70189	10,4%	65097	9,8%
Classe 4	33050	81,8%	32176	82,6%	30122	84,3%
Classe 5	29064	37,8%	27829	38,9%	24977	40,2%
Classe 6	4363	50,3%	3854	51,9%	2295	67,1%
Classe 7	20374	31,7%	20085	31,9%	17318	32,4%
Classe 8	165919	96,3%	165225	96,6%	158319	99,1%
Global	701600	73,5%	691827	74,0%	566646	74,6%

Tabela 29: Área de estudo A – Validação das amostras dos dados de treino criados a partir OSM: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos. Denota-se por S_DT_O0, S_DT_O1 e S_DT_O2 as amostras de DT_O0, DT_O1 e DT_O2, respectivamente.

	S_DT_O0		S_DT_O1		S_DT_O2	
	Nº Pixéis	% Corretos	Nº Pixéis	% Corretos	Nº Pixéis	% Corretos
Classe 1	89651	47,1%	88157	47,9%	35389	68,4%
Classe 2	120475	89,2%	109898	97,8%	94883	99,4%
Classe 3	22045	62,6%	19691	70,0%	16247	71,5%
Classe 4	336153	95,2%	321680	99,5%	302378	99,9%
Classe 5	371427	77,7%	359305	80,3%	344239	80,1%
Classe 6	3745	86,7%	3343	94,6%	2842	97,7%
Classe 7	0	-	0	-	0	-
Classe 8	9884	39,9%	9419	41,7%	3267	92,1%
Global	953380	81,8%	911493	85,5%	799245	89,3%

Tabela 30: Área de estudo B – Validação das amostras dos dados de treino criados a partir OSM: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos. Denota-se por S_DT_O0, S_DT_O1 e S_DT_O2 as amostras de DT_O0, DT_O1 e DT_O2, respectivamente.

Na Tabela 31, apresentam-se as exatidões globais dos MUCSs obtidos, observando que foi também com os dados de treino obtidos a partir do OSM com o critério O2 (Exclusividade) que se obteve MUCSs com melhor qualidade. Independentemente do critério utilizado, nenhum dos MUCSs criados a partir dos dados de treinos criados a partir do OSM apresenta melhores exatidões globais que os MUCSs obtidos com os dados de treino obtidos a partir da COS 2015.

	Área de Estudo A		Área de Estudo B	
	Classificação	Generalização	Classificação	Generalização
DT_O0_F2	61,0%	61,4%	67,9%	72,6%
DT_O1_F2	60,6%	60,6%	67,9%	70,1%
DT_O2_F2	61,7%	61,8%	70,1%	72,1%

Tabela 31: Exatidões globais das classificações, e respectivas generalizações, utilizando os dados de treino DT_O0_F2, DT_O1_F2 e DT_O2_F2.

No Anexo I são apresentados os MUCSs obtidos pela classificação de imagens de satélite e respectivas generalizações, utilizando como treino do classificador os dados DT_O0_F2, DT_O1_F2 e DT_O2_F2.

Nas figuras 33 e 34 apresentam-se para duas regiões escolhidas ao acaso nas áreas de estudo A e B, respetivamente, as classificações obtidas com os dados de treino extraídos diretamente do OSM.

Observando a Figura 34, a mesma região escolhida ao acaso para a área de estudo A analisada nas secções anteriores, em qualquer classificação e generalização, observa-se uma grande confusão na classificação para todas as classes podendo estar associado à baixa qualidade dos dados de treino para as classes 3, 5, 6 e 7 (Tabela 29). Para esta mesma região, temos melhores MUCSs obtidos pelos dados de treino criados pela COS 2015,

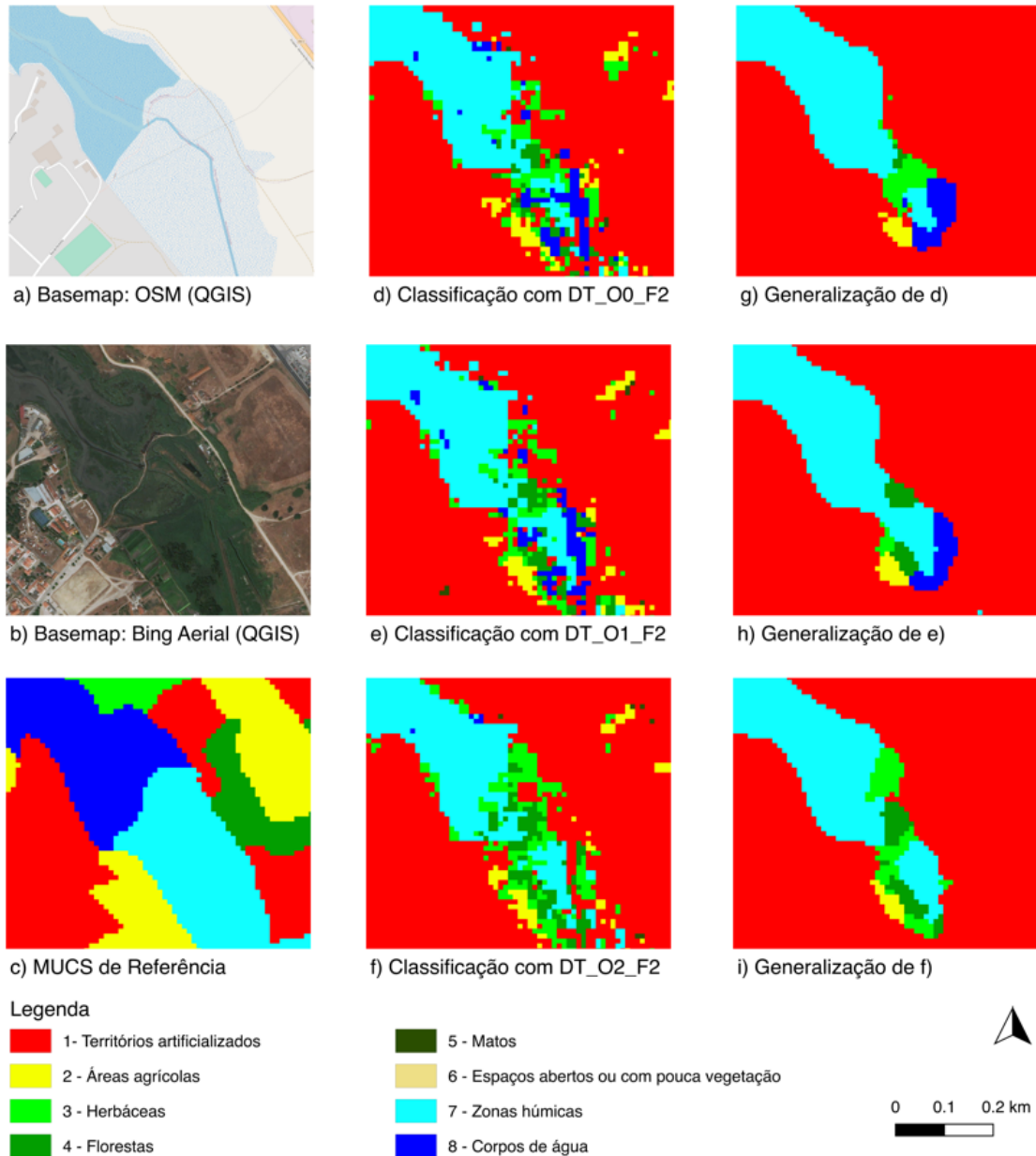


Figura 34: Região escolhida ao acaso na área de estudo A – MUCSs obtidos pelas classificações e respectivas generalizações a partir dos dados de treino criados a partir do OSM.

A partir da Figura 35, observa-se uma grande confusão na classificação entre as classes 5 e 6, sendo que esta última nem é considerada como existente para a região em questão. Apesar da qualidade elevada dos dados de treino para esta classe, observa-se que a quantidade destes dados é muito baixa quando comparado com os dados de treino para as restantes classes (Tabela 30). Ainda nesta figura, observa-se que para as classificações feitas existem vestígios de áreas classificadas como pertencentes à classe 1, correspondendo a estradas. Com a generalização feita para a produção da COS 2015 de forma a que a UMC seja 1ha, nos dados de treino criados a partir deste produtos ter-se-á poucos ou nenhuns dados de treino capazes de caracterizarem esta classe, não acontecendo

o mesmo com os dados de treino obtidos pelo OSM. Desta forma, tem-se uma área corretamente identificada mas que não está de acordo com o MUCS de referência que também é obtido por um produto com UMC de 1ha. Com a generalização feita, como estas áreas são pequenas são reclassificadas como pertencentes à classe que as rodeia, favorecendo a validação feita com o MUCS de referência.

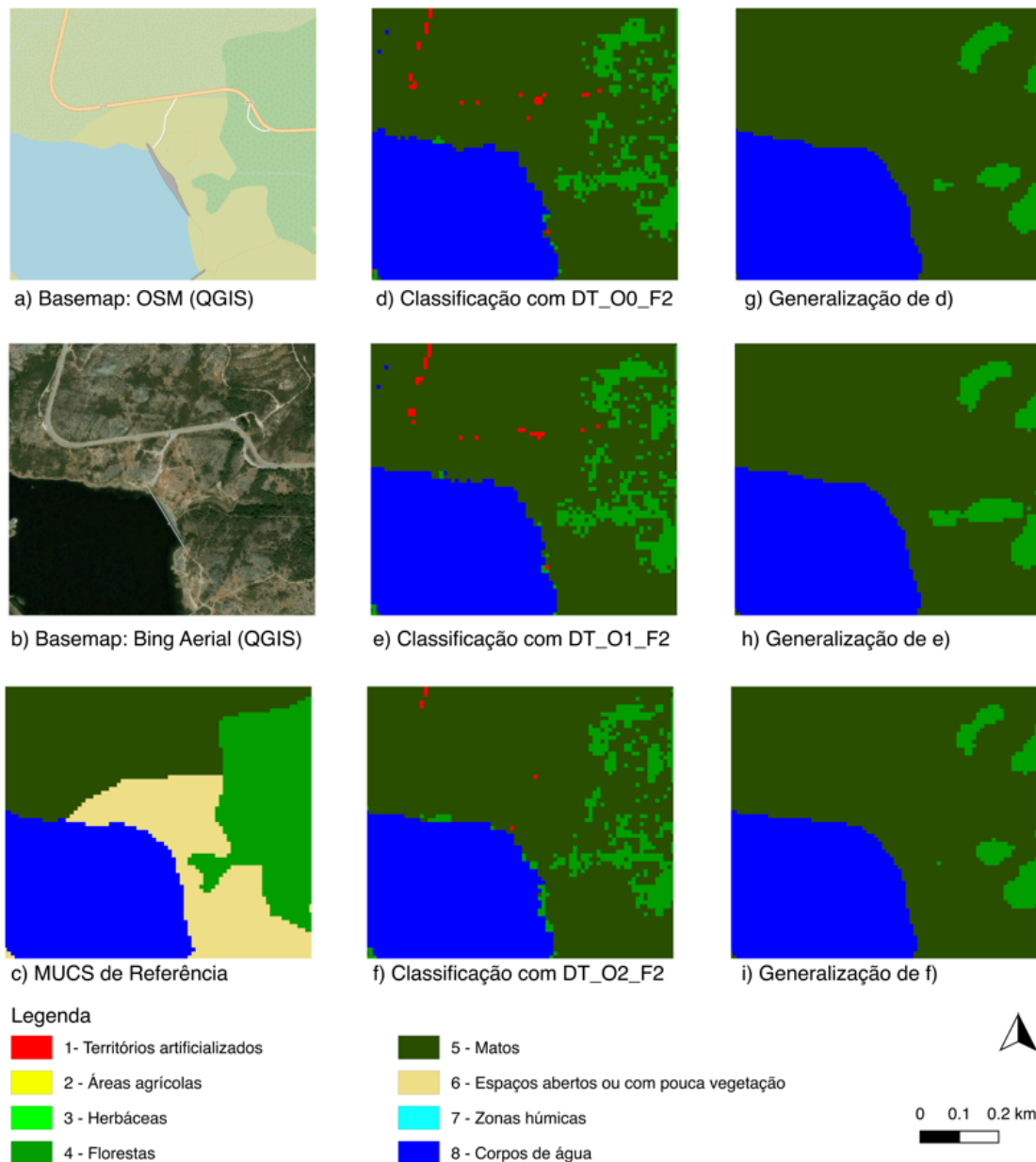


Figura 35: Região escolhida ao acaso na área de estudo B – MUCSs obtidos pelas classificações e respectivas generalizações a partir dos dados de treino criados a partir do OSM.

Nas tabelas 32 e 33 são apresentadas as exatidões do utilizador e do produtor, respetivamente, para os MUCSs obtidos na área de estudo A com dados de treino extraídos diretamente do OSM. A exatidão do utilizador dos MUCSs criados tende a aumentar com a generalização para todas as

classes expeto para as classes 1 e 3. Em contrapartida, a exatidão do produtor tende a diminuir para todas as classes exceto para as classes 1, 7 e 8. Com os dados de treino obtido a partir do OSM aplicando o critério O2 (Exclusividade), de forma geral, é aquele que apresenta melhores exatidões de utilizador e produtor. Quanto à exatidão do utilizador, esta apenas é melhor usando o critério O1 (Predominância) para as classes 5 e 7. No caso da exatidão do produtor, o critério O0 (Presença) apenas apresenta melhores resultados para a classe 1 e 6. Observa-se que para a área de estudo A, os erros de comissão são bastante significativos para as classes 3, 5 e 7 e os erros de omissão são bastante significativos para as classes 2, 3, 4, 5 e 6. Os maiores erros de omissão acontecem nas classes com menor concordância entre os dados de treino com a COS 2018, havendo assim menor concordância da classificação com as respetivas classes no mapa de referência.

	Exatidão do Utilizador												
	DT_O0_F2			DT_O1_F2			DT_O2_F2			O1-O0		O2-O1	
	C	G	G-C	C	G	G-C	C	G	G-C	C	G	C	G
Classe 1	45,0%	44,6%	-0,4%	44,7%	44,0%	-0,7%	45,8%	45,2%	-0,6%	-0,3%	-0,6%	0,9%	0,6%
Classe 2	79,5%	83,2%	3,8%	79,2%	82,8%	3,7%	80,8%	84,1%	3,3%	-0,3%	-0,4%	1,3%	0,9%
Classe 3	6,3%	5,7%	-0,7%	6,3%	5,5%	-0,7%	6,4%	5,7%	-0,6%	-0,1%	-0,1%	0,0%	0,1%
Classe 4	81,5%	82,3%	0,8%	81,4%	82,3%	0,8%	80,6%	81,1%	0,5%	-0,1%	0,0%	-0,9%	-1,2%
Classe 5	42,1%	46,6%	4,5%	42,2%	46,6%	4,4%	41,6%	46,1%	4,5%	0,1%	0,0%	-0,5%	-0,5%
Classe 6	58,8%	67,9%	9,1%	58,0%	67,2%	9,2%	68,4%	76,4%	7,9%	-0,8%	-0,7%	9,6%	8,4%
Classe 7	27,2%	28,2%	1,1%	27,2%	28,3%	1,1%	27,3%	28,4%	1,1%	0,1%	0,1%	0,1%	0,1%
Classe 8	98,9%	99,2%	0,2%	98,9%	99,1%	0,2%	99,4%	99,5%	0,0%	0,0%	0,0%	0,5%	0,3%

Tabela 32: Área de Estudo A – Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_O0_F2, DT_O1_F2 e DT_O2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

	Exatidão do Produtor												
	DT_O0_F2			DT_O1_F2			DT_O2_F2			O1-O0		O2-O1	
	C	G	G-C	C	G	G-C	C	G	G-C	C	G	C	G
Classe 1	98,0%	98,8%	0,8%	97,9%	98,8%	0,8%	97,8%	98,6%	0,8%	0,0%	0,0%	-0,1%	-0,1%
Classe 2	18,2%	16,1%	-2,1%	18,3%	16,2%	-2,1%	18,6%	16,6%	-2,0%	0,1%	0,1%	0,4%	0,5%
Classe 3	11,7%	9,5%	-2,2%	11,6%	9,3%	-2,3%	12,5%	10,3%	-2,3%	-0,1%	-0,2%	0,9%	0,7%
Classe 4	20,4%	20,1%	-0,3%	20,3%	20,0%	-0,3%	20,9%	20,4%	-0,5%	-0,1%	-0,1%	0,5%	0,3%
Classe 5	21,4%	19,2%	-2,2%	21,2%	19,0%	-2,3%	23,3%	21,0%	-2,3%	-0,2%	-0,2%	1,9%	1,8%
Classe 6	32,6%	31,8%	-0,7%	32,1%	31,3%	-0,8%	28,7%	27,5%	-1,2%	-0,5%	-0,6%	-3,9%	-4,3%
Classe 7	79,4%	81,9%	2,5%	79,3%	81,7%	2,5%	79,7%	82,2%	2,5%	-0,1%	-0,2%	0,3%	0,4%
Classe 8	69,5%	70,8%	1,3%	68,7%	69,1%	0,4%	70,6%	71,3%	0,7%	-0,8%	-1,6%	1,2%	0,5%

Tabela 33: Área de Estudo A – Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_O0_F2, DT_O1_F2 e DT_O2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

Quanto à área de estudo B, nas tabelas 34 e 35 são apresentadas as exatidões do utilizador e do produtor, respetivamente, dos MUCSs obtidos com dados de treino extraídos do OSM. A exatidão do utilizador dos MUCSs tende a aumentar com a generalização para todas as classes. Para esta área, quanto à exatidão do utilizador o uso do critério O2 (Exclusividade) mostrou ser o melhor, exceto para as classes 2, 4 e 5, onde com o critério O1 (Predominância) tem-se melhores resultados. A exatidão do produtor mostrou ser beneficiada apenas com o critério 2 (Exclusividade) para as classes 2, 4 e 5, sendo para as restantes beneficiadas com o critério 1 (Predominância). Observa-se que para a área de estudo A, os erros de comissão são bastante significativos para a classe 1 e os erros de omissão são bastante significativos para as classes 3 e 6. No caso da concordância ser superior 80% entre os dados de treino da classe 6 e o mapa de referência, a quantidade de dados de treino para esta classe são em menor número que qualquer outra classe. Desta forma, conclui-se que para esta classe os dados de treino poderão não ser suficientemente representativos, significando que dever-se-ia ter utilizado uma maior quantidade.

	Exatidão do Utilizador													
	DT_O0_F2			DT_O1_F2			DT_O2_F2			O1-O0		O2-O1		
	C	G	G-C	C	G	G-C	C	G	G-C	C	G	C	G	
Classe 1	28,8%	52,3%	23,4%	28,5%	32,1%	3,6%	43,3%	49,5%	6,2%	-0,3%	-20,2%	14,5%	-2,7%	
Classe 2	73,4%	73,9%	0,5%	73,4%	73,5%	0,1%	73,1%	73,0%	-0,1%	0,0%	-0,5%	-0,3%	-0,9%	
Classe 3	61,0%	93,6%	32,6%	61,2%	65,3%	4,1%	63,1%	73,0%	9,9%	0,2%	-28,3%	2,1%	-20,6%	
Classe 4	80,3%	78,8%	-1,5%	80,4%	82,1%	1,8%	80,0%	81,8%	1,8%	0,1%	3,3%	-0,3%	3,1%	
Classe 5	62,8%	66,7%	3,8%	63,0%	65,3%	2,3%	62,0%	64,3%	2,3%	0,1%	-1,4%	-0,9%	-2,3%	
Classe 6	95,0%	99,3%	4,3%	95,4%	98,8%	3,4%	97,4%	98,9%	1,4%	0,4%	-0,5%	2,4%	-0,4%	
Classe 8	55,1%	83,2%	28,1%	55,1%	76,9%	21,8%	81,2%	85,6%	4,4%	0,0%	-6,3%	26,2%	2,4%	

Tabela 34: Área de Estudo B – Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_O0_F2, DT_O1_F2 e DT_O2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

	Exatidão do Produtor													
	DT_O0_F2			DT_O1_F2			DT_O2_F2			O1-O0		O2-O1		
	C	G	G-C	C	G	G-C	C	G	G-C	C	G	C	G	
Classe 1	86,2%	79,3%	-6,8%	86,3%	88,2%	1,9%	76,1%	77,3%	1,2%	0,1%	8,8%	-10,1%	-2,0%	
Classe 2	69,7%	79,8%	10,1%	69,6%	72,7%	3,1%	77,4%	80,5%	3,2%	0,0%	-7,1%	7,7%	0,7%	
Classe 3	7,9%	4,1%	-3,8%	8,2%	4,4%	-3,8%	7,5%	4,1%	-3,4%	0,3%	0,3%	-0,4%	0,0%	
Classe 4	72,1%	78,9%	6,8%	72,3%	74,6%	2,3%	73,2%	75,1%	1,9%	0,1%	-4,3%	1,1%	-3,8%	
Classe 5	73,3%	73,3%	0,0%	73,1%	75,5%	2,5%	74,5%	76,6%	2,1%	-0,2%	2,3%	1,3%	3,4%	
Classe 6	1,1%	1,9%	0,8%	1,1%	0,5%	-0,7%	1,0%	0,4%	-0,6%	0,0%	-1,5%	-0,1%	-1,5%	
Classe 8	63,5%	54,6%	-8,9%	63,8%	58,9%	-4,9%	55,6%	52,2%	-3,4%	0,3%	4,3%	-7,9%	-2,4%	

Tabela 35: Área de Estudo B – Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_O0_F2, DT_O1_F2 e DT_O2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

4.2.4. Classificação com dados de treino extraídos da COS 2015 e do OSM

A partir dos dados de treino criados a partir do OSM e da COS 2015, podem-se retirar apenas os dados para cada classe que estão de acordo em ambos os produtos.

Na Tabela 36 e na Tabela 37, são apresentadas a percentagem de dados que são considerados como corretos quando comparados com o mapa de referência criado a partir da COS 2018 para as áreas de estudo A e B, respetivamente. Observa-se que de acordo com os critérios CO0 (Presença), CO1 (Predominância) e CO2 (Exclusividade), existe uma melhoria na qualidade dos dados de iteração para iteração, isto é, do critério CO0 (Presença) para o critério CO1 (Predominância) e do critério CO1 (Predominância) para o critério CO2 (Exclusividade).

	S_DT_CO0		S_DT_CO1		S_DT_CO2	
	Nº Pixéis	% Corretos	Nº Pixéis	% Corretos	Nº Pixéis	% Corretos
Classe 1	283491	99,3%	283171	99,4%	196649	99,4%
Classe 2	12207	96,9%	12139	97,5%	9922	97,5%
Classe 3	7775	84,3%	7640	85,8%	6331	85,8%
Classe 4	26541	98,7%	26430	99,1%	24866	99,2%
Classe 5	11529	89,5%	11434	90,2%	10312	90,3%
Classe 6	2020	95,5%	1979	97,5%	1428	98,2%
Classe 7	6224	99,8%	6211	100,0%	5298	100,0%
Classe 8	159668	100,0%	159590	100,0%	156704	100,0%
Global	509455	99,0%	508594	99,1%	411510	99,1%

Tabela 36: Área de estudo A – Validação das amostras dos dados de treino criados em simultâneo a partir OSM e da COS 2015: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos.

	S_DT_CO0		S_DT_CO1		S_DT_CO2	
	Nº Pixéis	% Corretos	Nº Pixéis	% Corretos	Nº Pixéis	% Corretos
Classe 1	41616	96,9%	40514	99,6%	21491	99,9%
Classe 2	95639	92,5%	88524	99,9%	77399	100,0%
Classe 3	10854	87,9%	9668	98,6%	8085	99,0%
Classe 4	319330	96,5%	308373	99,9%	290813	99,9%
Classe 5	282655	96,2%	273491	99,4%	259537	99,4%
Classe 6	3224	93,1%	2979	100,0%	2586	100,0%
Classe 8	3953	94,2%	3731	99,8%	2845	99,9%
Global	757271	95,7%	727280	99,7%	662756	99,7%

Tabela 37: Área de estudo B – Validação das amostras dos dados de treino criados em simultâneo a partir OSM e da COS 2015: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos.

Na Tabela 38 e na Tabela 39 são reunidos as exatidões globais dos dados de treino criados a partir da COS 2015, a partir do OSM e em simultâneo a partir de ambos os produtos. De forma geral, os dados de treino criados em simultâneo a partir de ambos os produtos apresentam sempre melhor qualidade que os dados de treino criados separadamente a partir de cada uma das fontes de dados. A única exceção à conclusão anterior ocorre para a classe 6 na área de estudo B, cuja exatidão dos dados de treino criados em simultâneo a partir das duas fontes é apenas melhor que os dados de treino criados a partir do OSM.

	C0	O0	CO0	C1	O1	CO1	C2	O2	CO2
% corretos									
Classe 1	93,2%	80,8%	99,3%	98,7%	81,0%	99,4%	98,7%	81,3%	99,4%
Classe 2	88,3%	67,9%	96,9%	95,4%	69,3%	97,5%	95,5%	72,4%	97,5%
Classe 3	55,6%	10,3%	84,3%	63,2%	10,4%	85,8%	63,3%	9,8%	85,8%
Classe 4	90,4%	81,8%	98,7%	95,9%	82,6%	99,1%	96,2%	84,3%	99,2%
Classe 5	79,7%	37,8%	89,5%	88,9%	38,9%	90,2%	89,2%	40,2%	90,3%
Classe 6	73,1%	50,3%	95,5%	87,5%	51,9%	97,5%	89,8%	67,1%	98,2%
Classe 7	93,3%	31,7%	99,8%	99,8%	31,9%	100,0%	99,9%	32,4%	100,0%
Classe 8	99,7%	96,3%	100,0%	100,0%	96,6%	100,0%	100,0%	99,1%	100,0%
Global	92,6%	73,5%	99,0%	96,6%	74,0%	99,1%	96,9%	74,6%	99,1%

Tabela 38: Área de estudo A – Validação das amostras dos dados de treino criados a partir OSM, da COS 2015 e em simultâneo por ambas as fontes de dados: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos. Para facilitar a leitura, os valores relativos ao uso em simultâneo da COS 2015 e do OSM encontram-se sombreados a verde quando são melhores que os valores relativos aos dados de treino criados exclusivamente a partir de cada fonte. Caso contrário, estes valores apresentam-se sombreados a vermelho.

	C0	O0	CO0	C1	O1	CO1	C2	O2	CO2
% corretos									
Classe 1	82,9%	47,1%	96,9%	97,4%	47,9%	99,6%	97,4%	68,4%	99,9%
Classe 2	87,4%	89,2%	92,5%	95,2%	97,8%	99,9%	95,4%	99,4%	100,0%
Classe 3	36,1%	62,6%	87,9%	42,0%	70,0%	98,6%	42,8%	71,5%	99,0%
Classe 4	90,6%	95,2%	96,5%	95,9%	99,5%	99,9%	95,9%	99,9%	99,9%
Classe 5	89,5%	77,7%	96,2%	94,4%	80,3%	99,4%	94,7%	80,1%	99,4%
Classe 6	93,8%	86,7%	93,1%	98,5%	94,6%	100,0%	99,0%	97,7%	100,0%
Classe 8	82,8%	39,9%	94,2%	99,4%	41,7%	99,8%	99,6%	92,1%	99,9%
Global	86,8%	81,8%	95,7%	93,1%	85,5%	99,7%	93,5%	89,3%	99,7%

Tabela 39: Área de estudo B – Validação das amostras dos dados de treino criados a partir OSM, da COS 2015 e em simultâneo por ambas as fontes de dados: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos. Para facilitar a leitura, os valores relativos ao uso em simultâneo da COS 2015 e do OSM encontram-se sombreados a verde quando são melhores que os valores relativos aos dados de treino criados exclusivamente a partir de cada fonte. Caso contrário, estes valores apresentam-se sombreados a vermelho.

Na Tabela 40, apresentam-se as exatidões globais dos MUCSs obtidos com os dados de treino extraídos em simultâneo da COS 2015 e do OSM. Observa-se que os MUCSs criados a partir destes dados de treino apresentam exatidões globais muito semelhantes entre si. Quando comparados com os MUCSs criados a partir unicamente da COS 2015 (Tabela 11, Tabela 12, Tabela 13 e Tabela 14) e unicamente pelos dados do OSM (Tabela 31), as exatidões globais obtidas encontram-se entre as exatidões globais dos MUCSs obtidos com os dados de treino extraídos do OSM e as exatidões globais dos MUCSs obtidos com os dados de treino extraídos da COS 2015.

	Área de Estudo A		Área de Estudo B	
	Classificação	Generalização	Classificação	Generalização
DT_CO0_F2	80,5%	80,5%	70,6%	72,6%
DT_CO1_F2	80,5%	80,5%	70,6%	72,6%
DT_CO2_F2	80,5%	80,5%	70,8%	72,7%

Tabela 40: Exatidões globais das classificações, e respetivas generalizações, utilizando os dados de treino DT_CO0_F2, DT_CO1_F2 e DT_CO2_F2.

Na Figura 36 e na Figura 37 são apresentadas os MUCSs obtidos para duas regiões contidas nas duas áreas de estudo usando os dados de treino extraídos a partir do OSM e da COS 2015. No Anexo J apresentam-se as classificações e respetivas generalizações efetuadas para ambas as áreas de estudo utilizando os dados de treino extraídos a partir do OSM e da COS 2015.

Observando a Figura 36, a mesma região escolhida ao acaso para a área de estudo A analisada nas secções anteriores, em qualquer classificação e generalização, observa-se uma grande confusão na classificação para todas as classes. Para esta mesma região, continua-se a ter melhores MUCSs obtidos pelos dados de treino criados unicamente a partir da COS 2015.

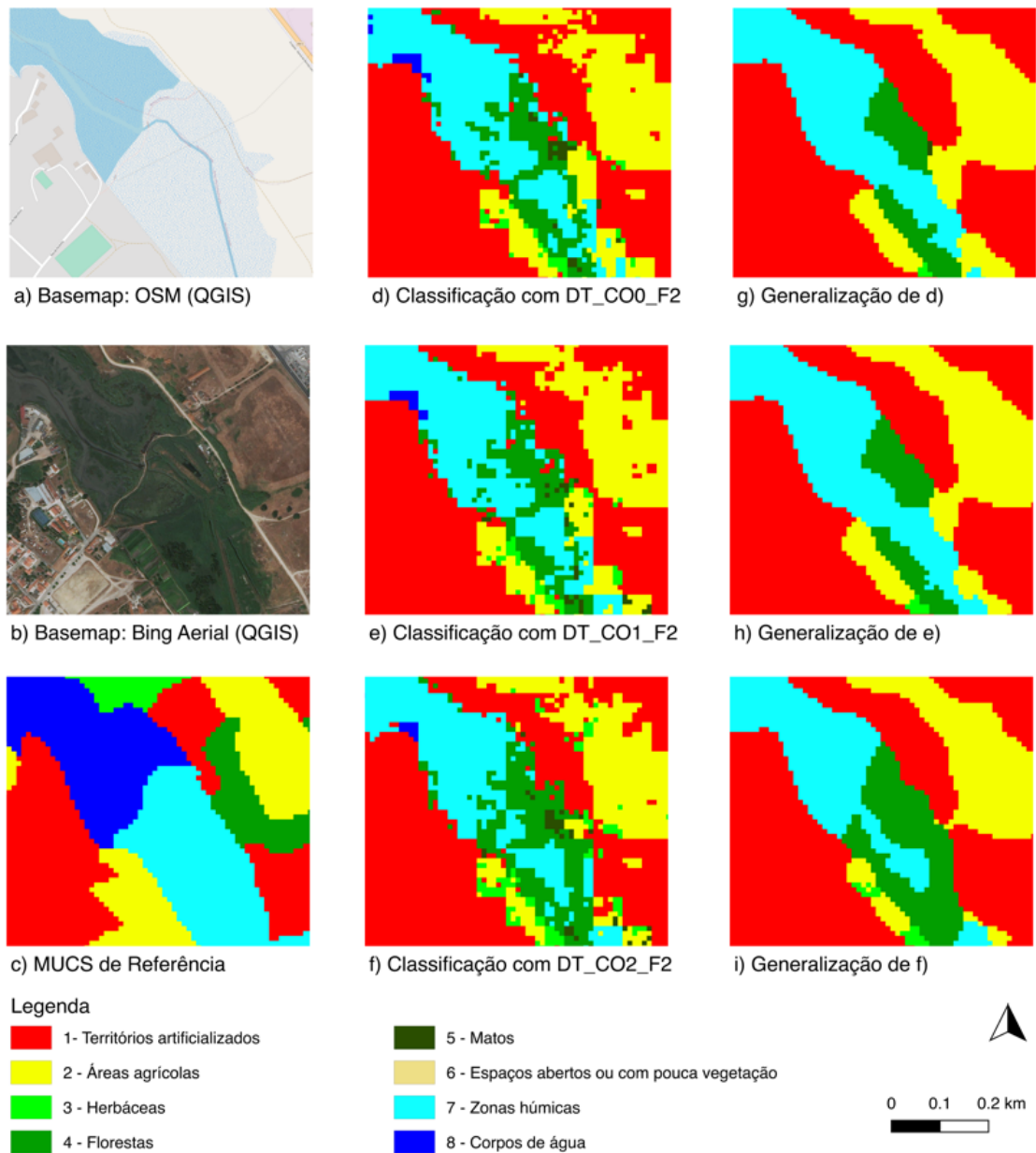


Figura 36: Região escolhida ao acaso na área de estudo A – MUCSs obtidos pelas classificações e respectivas generalizações a partir dos dados de treino criados a partir da COS 2015 e do OSM.

A partir da Figura 37, observa-se uma grande confusão na classificação entre as classes 5 e 6, tal e qual como tinha acontecido usando dados de treino obtidos exclusivamente pelo OSM, podendo estar associado à baixa quantidade de dados de treino usados para a classe 6.

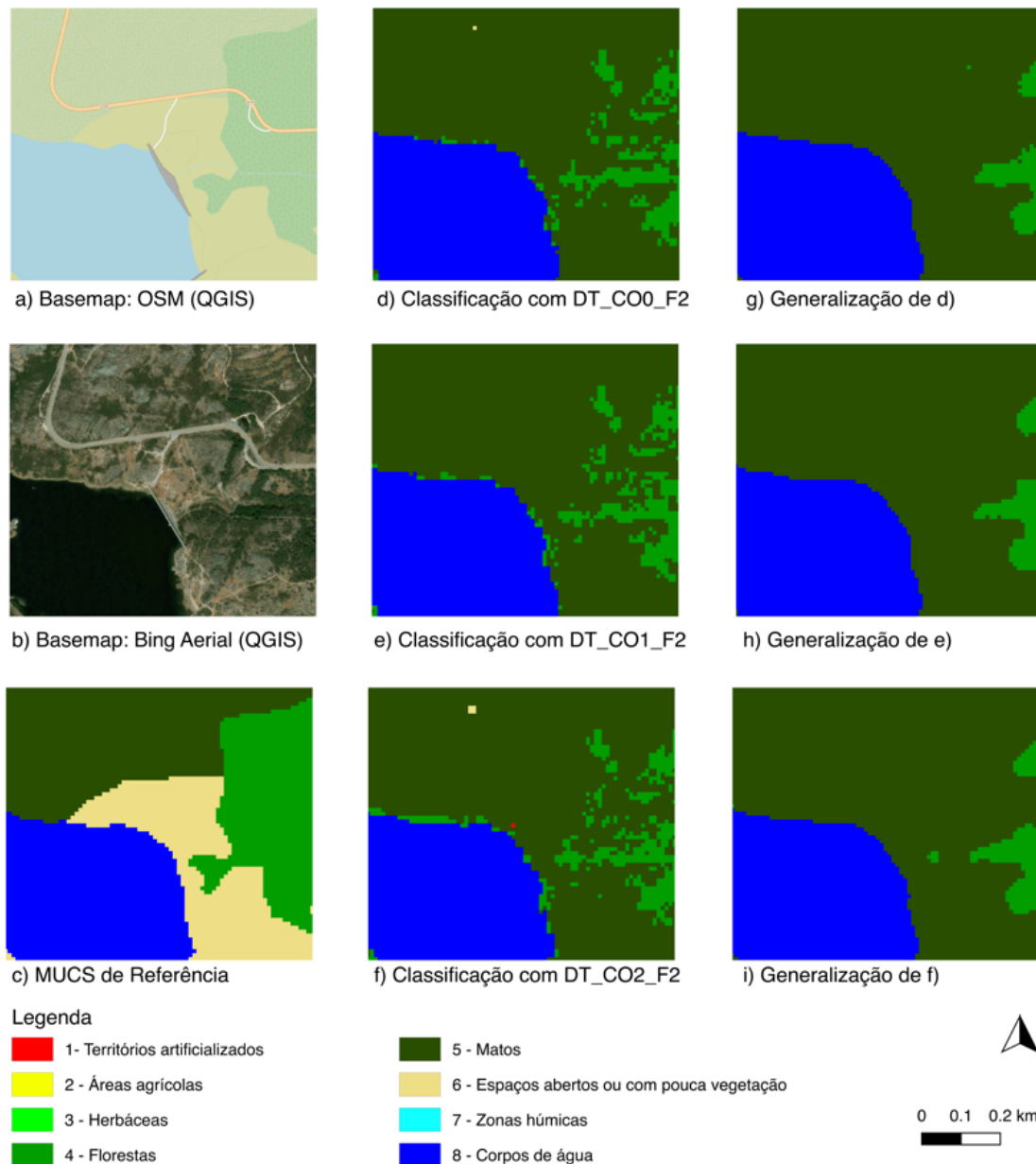


Figura 37: Região escolhida ao acaso na área de estudo B – MUCSs obtidos pelas classificações e respectivas generalizações a partir dos dados de treino criados a partir da COS 2015 e do OSM.

Para a área de estudo A, nas tabelas 41 e 42, com os dados de treino obtidos em simultâneo a partir da COS 2015 e pelo OSM, a exatidão do utilizador dos MUCSs tende a aumentar com a generalização para todas as classes expeto para a classe 1. Ainda quanto à generalização, a exatidão do produtor tende a diminuir para as classes 2, 3, 5 e 6 e a aumentar para as restantes. De forma geral, para esta área não existe um critério específico de obtenção dos dados de treino que melhor se adapte a todas as classes. Para a exatidão do utilizador, o critério CO0 (Presença) beneficia as classes 3, 4, 5 e 7 e o critério CO2 (Exclusividade) as classes restantes. Quanto à exatidão do produtor, o critério CO0 (Presença) apenas beneficia a classe 2, o critério CO1 (Predominância) beneficia as classes 1, 6, 7 e 8, e o critério CO2 (Exclusividade) as restantes.

Neste caso, a classe 3 também apresenta o maior erro de omissão e de comissão. Destaca-se ainda erros de omissão significativos para as classes 2, 5 e 6.

	Exatidão do Utilizador												
	DT_CO0_F2			DT_CO1_F2			DT_CO2_F2			CO1-CO0		CO2-CO1	
	C	G	G-C	C	G	G-C	C	G	G-C	C	G	C	G
Classe 1	62,0%	60,8%	-1,2%	62,0%	60,8%	-1,2%	62,4%	61,1%	-1,3%	0,0%	0,0%	0,3%	0,3%
Classe 2	73,4%	75,5%	2,1%	73,5%	75,7%	2,2%	73,6%	75,8%	2,1%	0,1%	0,2%	0,3%	0,3%
Classe 3	43,4%	53,5%	10,1%	43,0%	52,9%	9,9%	37,8%	46,8%	9,0%	-0,4%	-0,6%	-5,7%	-6,7%
Classe 4	85,0%	85,9%	1,0%	85,0%	85,9%	1,0%	84,3%	85,3%	0,9%	0,0%	0,0%	-0,6%	-0,7%
Classe 5	52,3%	59,3%	7,0%	52,2%	59,2%	7,0%	50,6%	57,6%	7,0%	-0,1%	-0,1%	-1,7%	-1,7%
Classe 6	54,4%	57,2%	2,8%	52,8%	55,3%	2,5%	59,2%	63,4%	4,2%	-1,7%	-1,9%	4,8%	6,1%
Classe 7	68,4%	71,6%	3,2%	67,6%	70,7%	3,1%	66,3%	69,8%	3,5%	-0,8%	-0,9%	-2,1%	-1,8%
Classe 8	99,3%	99,4%	0,0%	99,3%	99,4%	0,0%	99,4%	99,4%	0,0%	0,0%	0,0%	0,1%	0,1%

Tabela 41: Área de Estudo A – Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_CO0_F2, DT_CO1_F2 e DT_CO2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

	Exatidão do Produtor												
	DT_CO0_F2			DT_CO1_F2			DT_CO2_F2			CO1-CO0		CO2-CO1	
	C	G	G-C	C	G	G-C	C	G	G-C	C	G	C	G
Classe 1	97,6%	98,2%	0,7%	97,6%	98,2%	0,7%	97,4%	98,1%	0,7%	0,0%	0,0%	-0,2%	-0,1%
Classe 2	39,6%	38,7%	-0,9%	39,5%	38,6%	-0,9%	38,0%	37,2%	-0,8%	-0,1%	-0,1%	-1,6%	-1,5%
Classe 3	13,2%	11,0%	-2,3%	13,4%	11,2%	-2,3%	14,3%	12,0%	-2,3%	0,2%	0,2%	1,1%	1,0%
Classe 4	50,7%	50,8%	0,0%	50,8%	50,8%	0,1%	51,6%	51,7%	0,1%	0,0%	0,1%	0,8%	0,9%
Classe 5	30,3%	29,5%	-0,7%	30,3%	29,5%	-0,7%	31,8%	31,1%	-0,7%	0,0%	0,0%	1,5%	1,5%
Classe 6	40,0%	36,9%	-3,1%	40,0%	37,0%	-3,0%	37,2%	33,9%	-3,3%	0,0%	0,1%	-2,9%	-3,1%
Classe 7	72,8%	73,8%	1,1%	72,9%	74,0%	1,1%	71,6%	72,8%	1,1%	0,2%	0,2%	-1,1%	-1,0%
Classe 8	99,0%	99,1%	0,1%	99,0%	99,1%	0,1%	98,9%	99,0%	0,1%	0,0%	0,0%	-0,1%	-0,1%

Tabela 42: Área de Estudo A – Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_CO0_F2, DT_CO1_F2 e DT_CO2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

Sobre a área de estudo B, nas tabelas 43 e 44, com os dados de treino obtidos em simultâneo pelos dois produtos, observa-se que a exatidão do utilizador dos MUCSs tende a aumentar com a generalização para todas as classes exceto para a classe 2. Ainda quanto à generalização, a exatidão do produtor tende a diminuir para as classes 3, 6 e 8 e a aumentar para as restantes. De forma geral, também para esta área não existe um critério específico de obtenção dos dados de treino que melhor se adapte a todas as classes. Para a exatidão do utilizador, o critério CO0 (Presença) beneficia as

classes 2 e 6, enquanto que o critério CO1 (Presença) beneficia as classes 4 e 5 e o critério CO2 (Exclusividade) as classes restantes. Quanto à exatidão do produtor, o critério CO1 (Predominância) beneficia as classes 1, 3, 5 e 8 e o critério CO2 (Exclusividade) as restantes.

Para esta área destaca-se apenas a existência de erros de omissão significativos associados às classes 3 e 6.

Exatidão do Utilizador													
	DT_CO0_F2			DT_CO1_F2			DT_CO2_F2			CO1-CO0		CO2-CO1	
	C	G	G-C	C	G	G-C	C	G	G-C	C	G	C	G
Classe 1	47,7%	52,3%	4,6%	47,2%	51,8%	4,6%	53,8%	59,0%	5,2%	-0,5%	-0,4%	6,1%	6,7%
Classe 2	74,0%	73,9%	-0,1%	74,0%	73,9%	-0,1%	73,8%	73,5%	-0,2%	-0,1%	-0,1%	-0,3%	-0,4%
Classe 3	88,1%	93,6%	5,5%	88,1%	93,6%	5,5%	89,6%	94,4%	4,8%	-0,1%	0,0%	1,4%	0,7%
Classe 4	77,1%	78,8%	1,7%	77,0%	78,7%	1,7%	76,8%	78,5%	1,7%	0,0%	-0,1%	-0,2%	-0,3%
Classe 5	63,8%	66,7%	2,9%	64,0%	66,8%	2,9%	63,3%	66,2%	2,9%	0,2%	0,2%	-0,5%	-0,5%
Classe 6	96,4%	99,3%	2,8%	96,2%	98,8%	2,7%	96,1%	99,3%	3,2%	-0,3%	-0,5%	-0,3%	0,0%
Classe 8	77,2%	83,2%	6,0%	77,0%	82,8%	5,8%	82,7%	86,2%	3,5%	-0,2%	-0,4%	5,5%	3,0%

Tabela 43: Área de Estudo B – Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_CO0_F2, DT_CO1_F2 e DT_CO2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

Exatidão do Produtor													
	DT_CO0_F2			DT_CO1_F2			DT_CO2_F2			CO1-CO0		CO2-CO1	
	C	G	G-C	C	G	G-C	C	G	G-C	C	G	C	G
Classe 1	78,5%	79,3%	0,9%	78,7%	79,6%	0,8%	70,9%	71,6%	0,7%	0,3%	0,2%	-7,5%	-7,7%
Classe 2	77,0%	79,8%	2,8%	77,0%	79,9%	2,9%	78,7%	81,5%	2,8%	0,0%	0,1%	1,7%	1,7%
Classe 3	6,5%	4,1%	-2,4%	6,6%	4,1%	-2,5%	6,0%	3,7%	-2,3%	0,1%	0,0%	-0,5%	-0,4%
Classe 4	76,8%	78,9%	2,1%	77,0%	79,1%	2,1%	76,9%	78,9%	2,0%	0,2%	0,2%	0,1%	0,0%
Classe 5	71,3%	73,3%	1,9%	71,1%	73,0%	1,9%	71,5%	73,3%	1,8%	-0,2%	-0,3%	0,2%	0,0%
Classe 6	3,0%	1,9%	-1,0%	3,1%	2,0%	-1,1%	3,0%	1,9%	-1,2%	0,1%	0,0%	0,1%	-0,1%
Classe 8	59,8%	54,6%	-5,2%	60,1%	54,8%	-5,3%	54,5%	51,7%	-2,8%	0,3%	0,2%	-5,3%	-2,9%

Tabela 44: Área de Estudo B – Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_CO0_F2, DT_CO1_F2 e DT_CO2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

4.2.5. Classificação com dados de treino extraídos do UA 2012

Usando o UA 2012 para criar dados de treino, apenas ter-se-á dados de treino para a área de estudo A, uma vez que nenhuma das cidades que engloba a área de estudo B tem o número mínimo de habitantes estabelecido para a produção do UA. Na Figura 38 é apresentada a área coberta pelas áreas de treino na área de estudo A, observando-se que apenas a zona do oceano não é coberta por completo. Na figura, a área coberta pelas áreas de treino DT_U0 e DT_U1 é constituída pelas áreas a amarelo e pelas áreas a azul. No caso da área coberta pelas áreas de treino DT_U2 esta é apresentada a amarelo.

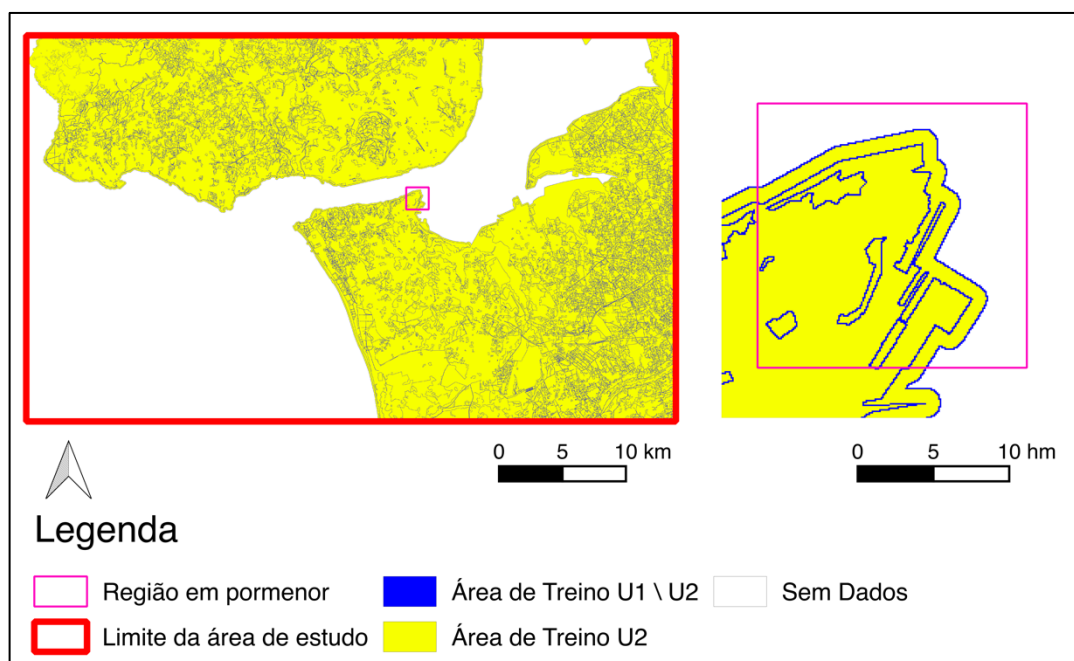


Figura 38: Área de estudo A – Áreas cobertas pelos dados de treino obtidos a partir do UA 2012 (azul – área coberta por DT_U0 e DT_U1; amarelo – área pertencente a DT_U0, DT_U1 e DT_U2).

Nas Tabela 45 são apresentadas as comparações dos dados de treino criados para a área de estudo A com o mapa de referência. Em geral, os melhores dados de treino obtidos com o OSM são criados utilizando o critério O2 (Exclusividade), sendo que as menores concordâncias ocorrem para as classes 3 e 7.

	S_DT_U0		S_DT_U1		S_DT_U2	
	Nº Pixéis	% Corretos	Nº Pixéis	% Corretos	Nº Pixéis	% Corretos
Classe 1	452054	78,8%	417103	83,0%	387376	85,8%
Classe 2	133271	52,4%	117068	54,3%	100615	55,7%
Classe 3	242587	29,9%	218303	31,4%	193811	32,6%
Classe 4	112162	89,8%	105619	91,4%	98733	92,5%
Classe 6	4829	65,2%	4154	70,8%	3559	75,0%
Classe 7	11741	39,9%	10930	41,0%	10117	41,8%
Classe 8	46039	88,1%	44071	90,5%	40275	92,0%
Global	1002683	64,6%	917248	67,9%	834486	70,3%

Tabela 45: Área de estudo A – Validação das amostras dos dados de treino criados a partir do UA 2012: números de pixéis que constituem os dados de treino e percentagem de pixéis corretos.

Denota-se por S_DT_U0, S_DT_U1 e S_DT_U2 as amostras de DT_U0, DT_U1 e DT_U2, respetivamente.

Observando a Figura 39, tem-se a mesma região escolhida ao acaso para a área de estudo A analisada nas secções anteriores, com as classificações e generalizações efetuadas. Observa-se uma grande confusão na classificação entre a classe 3 e as restantes, entre as classes 1 e 2 e entre as classes 7 e 8. Para esta mesma região, continua-se a ter melhores MUCSs obtidos pelos dados de treino criados unicamente a partir da COS 2015. No Anexo K tem-se as classificações e generalizações feitas para a área de estudo A utilizando os dados de treino extraídos do UA 2012.

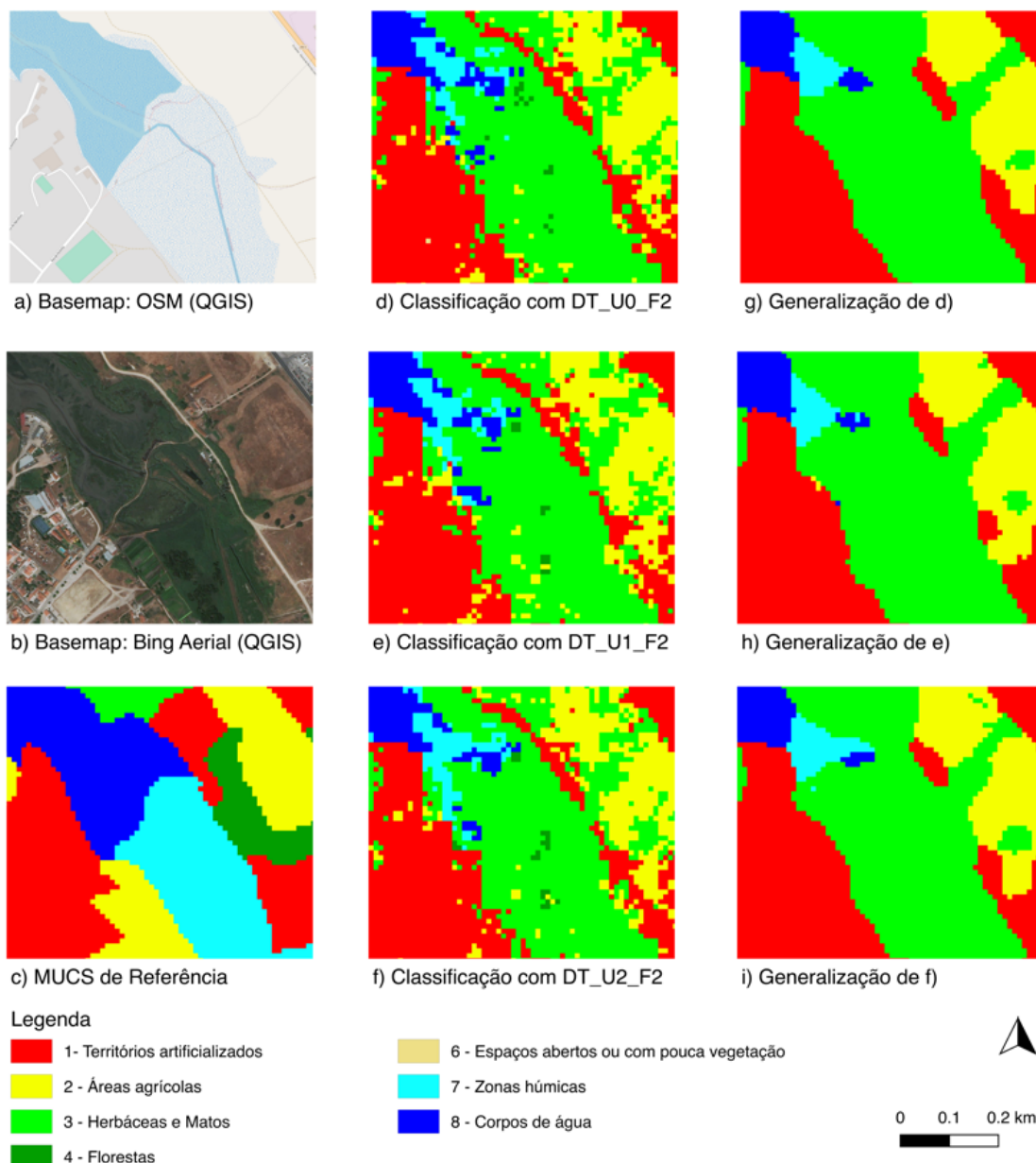


Figura 39: Região escolhida ao acaso na área de estudo A – MUCSs obtidos pelas classificações e respetivas generalizações a partir dos dados de treino criados a partir do UA 2012.

Na Tabela 46, apresenta-se as exatidões globais das classificações e generalizações a partir dos dados de treino extraídos do UA 2012. Observa-se que os MUCSs criados apresentam exatidões globais muito semelhantes entre si, não apresentando melhores resultados que os obtidos com a COS 2015. É de observar que neste caso prático, tem-se uma classe a menos, uma vez que 2 das classes consideradas anteriormente são agregadas numa só. Desta forma, é de esperar que se a classificação efetuada com os dados de treino extraídos da COS 2015 fosse adaptada a este conjunto de classes a sua exatidão global fosse superior.

	Área de Estudo A	
	Classificação	Generalização
DT_U0_F2	80,7%	81,1%
DT_U1_F2	80,7%	81,2%
DT_U2_F2	80,7%	81,1%

Tabela 46: Exatidões globais das classificações, e respectivas generalizações, utilizando os dados de treino DT_U0_F2, DT_U1_F2 e DT_U2_F2.

Nas tabelas 47 e 48 são apresentadas as exatidões do utilizador e do produtor, respetivamente, para os MUCSs obtidos com os dados de treino extraídos do UA 2012. A exatidão do utilizador dos MUCSs tende a aumentar com a generalização para todas as classes expeto para a classe 1 e a diminuir a exatidão do produtor das classes 2 e 6. De forma geral, para esta área não existe um critério específico de obtenção dos dados de treino que melhor se adapte a todas as classes. Para a exatidão do utilizador, o critério 0 (Presença) apenas beneficia as classes 2 e 3, o critério 1 (Predominância) a classe 7 e o critério 2 (Exclusividade) as classes restantes. Quanto à exatidão do produtor, o critério 0 (Presença) beneficia as classes 1 e 4, o critério 1 (Predominância) beneficia as classes 6 e 8, e o critério 2 (Exclusividade) as restantes.

Os maiores erros de comissão pertencem às classes 3 e 7, sendo as mesmas classes que têm dados de treino com menor concordância com o mapa de referência. Quanto aos erros de omissão, estes ocorrem com maior impacto para quase todas as classes, exceto para as classes 1 e 8.

	Exatidão do Utilizador												
	DT_U0_F2			DT_U1_F2			DT_U2_F2			U1-U0		U2-U1	
	C	G	G-C	C	G	G-C	C	G	G-C	C	G	C	G
Classe 1	77,7%	77,0%	-0,7%	78,2%	77,6%	-0,7%	78,2%	77,4%	-0,8%	0,5%	0,6%	0,5%	0,4%
Classe 2	68,2%	75,0%	6,7%	67,6%	74,5%	7,0%	68,6%	75,0%	6,4%	-0,7%	-0,4%	0,4%	0,1%
Classe 3	30,3%	30,9%	0,5%	30,2%	30,8%	0,6%	30,1%	30,6%	0,5%	-0,1%	-0,1%	-0,2%	-0,3%
Classe 4	91,8%	93,0%	1,2%	91,8%	93,0%	1,2%	91,8%	92,9%	1,1%	0,0%	0,0%	0,0%	-0,1%
Classe 6	80,2%	85,0%	4,8%	80,1%	85,2%	5,1%	81,0%	85,5%	4,5%	-0,1%	0,2%	0,8%	0,5%
Classe 7	41,5%	41,9%	0,5%	41,5%	42,1%	0,6%	41,4%	41,8%	0,4%	0,1%	0,2%	-0,1%	-0,1%
Classe 8	99,4%	99,3%	0,0%	99,4%	99,3%	0,0%	99,4%	99,4%	0,0%	0,0%	0,0%	0,0%	0,0%

Tabela 47: Área de Estudo A – Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_U0_F2, DT_U1_F2 e DT_U2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

Exatidão do Produtor													
	DT_U0_F2			DT_U1_F2			DT_U2_F2			U1-U0		U2-U1	
	C	G	G-C	C	G	G-C	C	G	G-C	C	G	C	G
Classe 1	93,8%	95,4%	1,6%	93,4%	95,1%	1,7%	93,6%	95,1%	1,5%	-0,4%	-0,2%	-0,1%	-0,3%
Classe 2	37,0%	36,4%	-0,6%	37,6%	37,1%	-0,5%	37,2%	36,6%	-0,6%	0,6%	0,7%	0,2%	0,2%
Classe 3	57,9%	58,0%	0,2%	58,4%	58,6%	0,2%	58,5%	58,6%	0,1%	0,5%	0,6%	0,6%	0,5%
Classe 4	44,9%	45,0%	0,2%	44,8%	44,9%	0,1%	44,9%	44,9%	0,0%	-0,1%	-0,1%	0,0%	-0,2%
Classe 6	35,4%	34,0%	-1,4%	35,7%	34,1%	-1,6%	34,7%	33,1%	-1,6%	0,3%	0,0%	-0,7%	-0,9%
Classe 7	35,4%	35,9%	0,5%	35,6%	36,3%	0,7%	36,1%	36,5%	0,4%	0,1%	0,3%	0,7%	0,6%
Classe 8	98,8%	98,9%	0,0%	98,8%	98,9%	0,0%	98,8%	98,8%	0,0%	0,0%	0,0%	0,0%	0,0%

Tabela 48: Área de Estudo A – Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_U0_F2, DT_U1_F2 e DT_U2_F2. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

4.2.6. Classificação com dados de treino extraídos da COS 2015 Utilizando os Vizinhos como *Features*

A partir das áreas de treino DT_C1 podem-se associar o conjunto de features F1 ou F1-4 de forma a comparar o uso da informação dos pixéis vizinhos ser vantajosa ou não no seu uso. Na secção 4.2.1 mostrou-se que o uso das *features* F2 era mais vantajoso que F1. Usando também a informação que caracteriza os 4 vizinhos em torno de um pixel, o número de *features* aumenta com um fator multiplicativo 5, significando que usando as *features* F1, em vez de se ter 12 *features* ter-se-á 60, e com F2 em vez de se ter 36 *features* ter-se-á 180. Por este motivo, o esforço computacional é muito maior, e por limitações computacionais apenas se pretende comparar se o uso da informação dos pixéis vizinhos é vantajosa ou não usando F1. Assim, tem-se os dados de treino DT_C1_F1 e DT_C1_F1-4.

A partir da Figura 40 e da Figura 41, têm-se as classificações para as mesmas regiões escolhidas ao acaso analisadas nos resultados anteriores, para as áreas de estudo A e B, respetivamente. Observa-se que nas classificações utilizando DT_C1_F1 existe muita confusão na classificação. Para a área de estudo A, com DT_C1_F1-4, por exemplo, observa-se melhorias na distinção entre as classes 1 e 2 (lado direito da região em pormenor). No caso da área de estudo B, o uso adicional das *features* que caracterizam os quatro pixéis vizinhos, parece existir melhorias na distinção da classe 4 com as restantes. No Anexo L têm-se as classificações e generalizações feitas para ambas as áreas de estudo com os dados de treino criados.

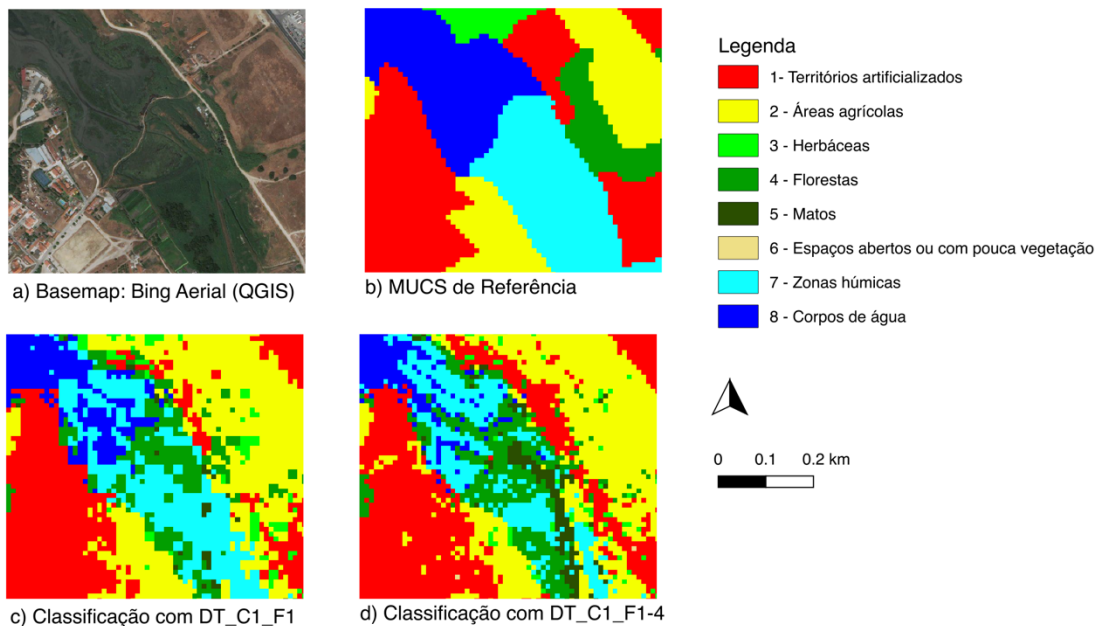


Figura 40: Região escolhida ao acaso na área de estudo A – MUCSs obtidos pelas classificações a partir dos dados de treino criados a partir do COS 2015 utilizando como *features* as bandas de 10m no pixel e nos pixéis vizinhos (Norte, Sul, Este e Oeste).

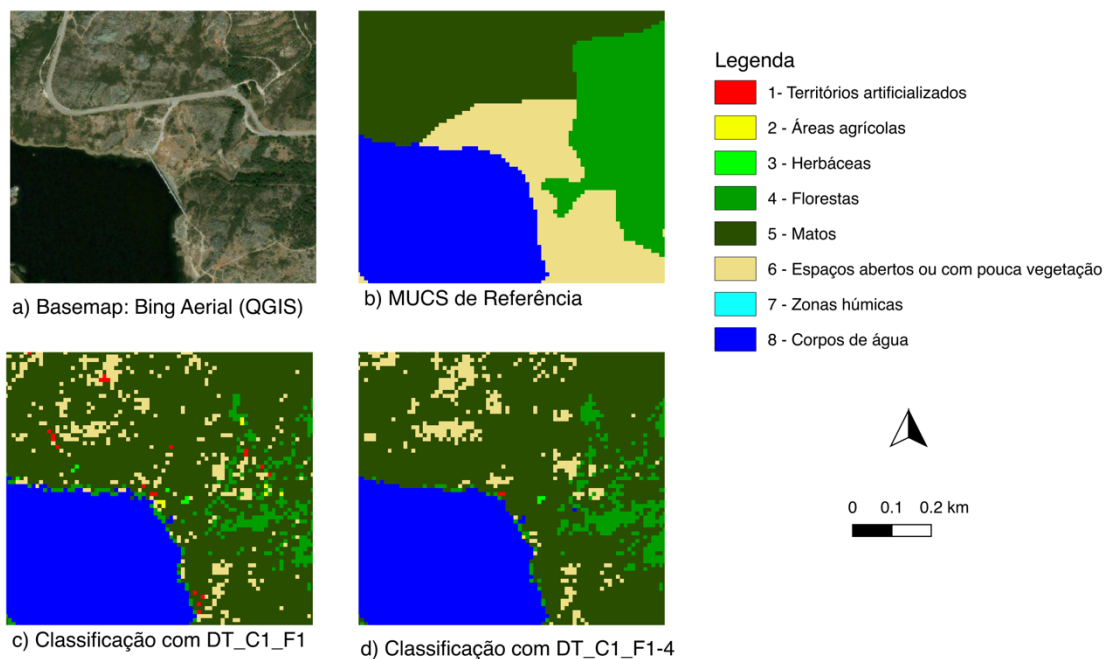


Figura 41: Região escolhida ao acaso na área de estudo B – MUCSs obtidos pelas classificações a partir dos dados de treino criados a partir do COS 2015 utilizando como *features* as bandas de 10m no pixel e nos pixéis vizinhos (Norte, Sul, Este e Oeste).

Na Tabela 48 são apresentadas as exatidões globais dos MUCSs criados para ambas as áreas de estudo a partir das classificações utilizando os dados de treino mencionados anteriormente. Observa-se que para ambas as áreas de estudo, o uso adicional das *features* que caracterizam os pixéis vizinhos é vantajosa, tendo-se aumentos das exatidões globais. Tal e qual como anteriormente, o processo de generalização usado cria MUCSs com exatidões globais maiores.

	Lisboa		Serra da Estrela	
	Classificação	Generalização	Classificação	Generalização
DT_C1_F1	84,9%	87,2%	70,7%	75,0%
DT_C1_F1-4	86,5%	87,7%	73,2%	75,4%

Tabela 49: Exatidões globais das classificações e generalizações utilizando os dados de treino DT_C1_F1 e DT_C1_F1-4.

Nas tabelas 50 e 51, apresentam-se as exatidões do utilizador e do produtor, respetivamente, para a área de estudo A. Para ambas as exatidões, estas aumentaram usando DT_C1_F1-4 em vez de DT_C1_F1, exceto para a exatidão do utilizador da classe 8, onde o decréscimo não foi significativo. O erro de comissão maior para esta área ocorreu com a classe 3, tendo sido bastante melhorada com o processo de generalização. No caso dos erros de omissão, estes ocorres com maior impacto nas classes 3, 5 e 6.

	Exatidão do Utilizador							
	DT_C1_F1			DT_C1_F1-4			DT_C1_F1-4 - DT_C1_F1	
	C	G	G-C	C	G	G-C	C	G
Classe 1	78,2%	77,7%	-0,5%	80,7%	79,7%	-1,1%	2,5%	1,9%
Classe 2	62,8%	72,2%	9,5%	66,2%	71,9%	5,7%	3,5%	-0,3%
Classe 3	41,3%	66,2%	25,0%	46,8%	65,3%	18,5%	5,5%	-1,0%
Classe 4	75,7%	79,4%	3,7%	77,8%	79,4%	1,6%	2,0%	-0,1%
Classe 5	62,7%	84,3%	21,5%	70,6%	81,6%	11,0%	7,8%	-2,7%
Classe 6	83,7%	94,9%	11,2%	87,1%	94,2%	7,1%	3,4%	-0,7%
Classe 7	84,9%	92,0%	7,1%	89,1%	91,7%	2,6%	4,2%	-0,3%
Classe 8	99,5%	99,5%	0,0%	99,4%	99,5%	0,1%	-0,1%	0,0%

Tabela 50: Área de Estudo A – Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_C1_F1 e DT_C1_F1-4. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

	Exatidão do Produtor							
	DT_C1_F1			DT_C1_F1-4			DT_C1_F1-4 - DT_C1_F1	
	C	G	G-C	C	G	G-C	C	G
Classe 1	88,4%	94,6%	6,2%	90,8%	94,1%	3,3%	2,3%	-0,5%
Classe 2	61,4%	66,5%	5,1%	65,9%	69,8%	3,9%	4,5%	3,3%
Classe 3	25,7%	19,4%	-6,3%	27,6%	21,5%	-6,1%	1,9%	2,1%
Classe 4	80,8%	85,7%	4,9%	83,4%	85,9%	2,6%	2,6%	0,3%
Classe 5	29,2%	24,8%	-4,5%	31,7%	27,7%	-4,1%	2,5%	2,9%
Classe 6	39,4%	32,9%	-6,4%	44,8%	38,0%	-6,8%	5,4%	5,1%
Classe 7	66,8%	70,6%	3,8%	69,8%	72,0%	2,1%	3,0%	1,4%
Classe 8	99,4%	99,6%	0,2%	99,5%	99,6%	0,1%	0,1%	0,1%

Tabela 51: Área de Estudo A – Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_C1_F1 e DT_C1_F1-4. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

No caso da área de estudo B, as exatidões do utilizador e do produtor são apresentadas nas tabelas 52 e 53, respetivamente. Verifica-se que as exatidões do utilizador e do produtor aumentaram usando DT_C1_F1-4 em vez de DT_C1_F1, exceto para a exatidão do utilizador da classe 8, onde o decréscimo bastante significativo, significando que poderão existir pequenos lagos que não aparecem na COS 2018 por causa da UMC estabelecida. O maior erro de comissão acontece para a classe 3 e de omissão para as classes 3 e 6.

	Exatidão do Utilizador							
	DT_C1_F1			DT_C1_F1-4			DT_C1_F1-4 - DT_C1_F1	
	C	G	G-C	C	G	G-C	C	G
Classe 1	72,4%	89,4%	17,0%	75,9%	86,4%	10,4%	3,5%	-3,0%
Classe 2	73,1%	73,5%	0,4%	75,2%	74,7%	-0,5%	2,1%	1,3%
Classe 3	44,8%	88,8%	44,0%	46,9%	89,3%	42,4%	2,1%	0,5%
Classe 4	73,4%	76,9%	3,5%	75,3%	77,0%	1,7%	1,9%	0,1%
Classe 5	64,8%	72,0%	7,3%	68,4%	72,1%	3,7%	3,6%	0,1%
Classe 6	70,5%	85,5%	15,0%	77,2%	85,3%	8,1%	6,7%	-0,2%
Classe 8	90,6%	93,5%	2,8%	61,9%	93,4%	31,6%	-28,8%	0,0%

Tabela 52: Área de Estudo B – Exatidão do utilizador para os MUCSs obtidos a partir dos dados de treino DT_C1_F1 e DT_C1_F1-4. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

	Exatidão do Produtor							
	DT_C1_F1			DT_C1_F1-4			DT_C1_F1-4 - DT_C1_F1	
	C	G	G-C	C	G	G-C	C	G
Classe 1	48,0%	47,4%	-0,7%	57,0%	54,9%	-2,1%	9,0%	7,5%
Classe 2	80,3%	86,6%	6,3%	81,7%	85,5%	3,8%	1,3%	-1,1%
Classe 3	10,3%	2,4%	-7,9%	10,2%	2,6%	-7,6%	0,0%	0,2%
Classe 4	78,0%	82,7%	4,7%	80,6%	83,0%	2,4%	2,6%	0,3%
Classe 5	64,9%	69,9%	5,0%	67,8%	70,8%	2,9%	3,0%	0,9%
Classe 6	41,4%	39,0%	-2,4%	43,0%	39,9%	-3,1%	1,6%	0,9%
Classe 8	54,6%	49,7%	-4,9%	55,5%	50,5%	-5,1%	0,9%	0,8%

Tabela 53: Área de Estudo B – Exatidão do produtor para os MUCSs obtidos a partir dos dados de treino DT_C1_F1 e DT_C1_F1-4. Para facilitar a leitura os valores positivos encontram-se a verde e os negativos a vermelho. No caso dos valores nulos, estes encontram-se a verde ou vermelho quando as diferenças obtidas utilizando um maior número de casas decimais são positivas ou negativas, respetivamente. (C – Classificação; G – Generalização)

Na Tabela 54 é apresentado a percentagem de pixéis classificados corretamente apenas na área de mudança para ambas as áreas de estudo. O uso adicional das *features* que caracterizam os 4 pixéis vizinhos mostrou aumentar a percentagem de área corretamente classificada onde ocorreu mudança da COS 2015 para a COS 2018.

	Área de Mudança			
	Área de Estudo A		Área de Estudo B	
	Classificação	Generalização	Classificação	Generalização
DT_C1_F1	37,3%	42,6%	44,0%	48,9%
DT_C1_F1-4	39,5%	44,2%	45,0%	49,0%

Tabela 54: Percentagem de área corretamente classificada nas áreas onde ocorreram mudança de ocupação e uso do solo, da COS 2015 para COS 2018.

Conclusão

De forma geral, o uso do conjunto de *features* F2 mostrou ser mais vantajoso em relação ao uso de F1. Com o conjunto F2, utilizam-se dados de treino também caracterizados por um maior número de respostas espectrais com diferentes comprimentos de onda, principalmente na zona do espectro dos infravermelhos. Quanto ao uso do conjunto de *features* F3, este mostrou ser menos vantajoso em relação ao uso de F2 quando usando dados de treino extraídos da COS 2015, provavelmente porque terá diminuído a heterogeneidade dos dados, ou seja, haverá dados de treino pertencentes a classes diferentes com características muito semelhantes nas bandas adicionadas. No entanto, quanto maior o número de *features* maior é a exigência computacional. Logo, o uso de F3 é mais exigente computacionalmente do que F2 e F2 é mais exigente que F1.

Com os dados de treino criados diretamente com a COS 2015, usando os critérios C0 (Presença), C1 (Predominância) e C2 (Exclusividade), são sucessivamente eliminadas áreas de treino. Na primeira iteração, os pixels que com o critério C0 eram usados como treino para mais que uma classe, no critério C1 são usados como treino apenas para a classe majoritária. Conclui-se assim que o mesmo pixel pertencer aos dados de treino de mais do que uma classe provoca confusão no processo de classificação. Na segunda iteração, do critério C1 para C2, são eliminados todos os pixels que não estão totalmente contidos dos polígonos provenientes da COS 2015. Com esta iteração, para a COS 2015 todos os pixels mistos são eliminados, o que diminuiu em geral as exatidões para as duas áreas de estudo. Isto pode ser justificado pelo facto de a UMC da COS 2018 ser 1 ha, o que implica que algumas classes poderão ser mais heterogêneas, como zonas urbanas de baixa densidade e zonas de matos com mistura de solo nu e herbáceas. Logo, ao se excluírem pixels mistos dos dados de treino estas zonas podem ficar mais mal caracterizadas. Assim, usando a COS 2015 como fonte de dados de treino e a COS 2018 como dados de referência, conclui-se que a existência de pixels mistos nos dados de treino poderá ser importante para algumas classes.

Usando o mapa de referência criado a partir da COS 2018, todos os MUCS criados com as generalizações das classificações mostraram ter melhor qualidade em relação os MUCS criados a partir das classificações das imagens de satélite. Um dos possíveis motivos do sucedido está na UMC da COS 2018 ser 1ha, dado que com as generalizações dos MUCS uma grande parte das áreas mais pequenas desaparecem.

A utilização das filtragens nas áreas de treino obtidas a partir da COS 2015 em função dos índices radiométricos não mostrou levar a melhorias na classificação das imagens de satélite. Com o aumento da separabilidade, aumenta-se a homogeneidade dos dados correspondentes a cada

classe, significando que, dada a UMC da COS 2018 removeram-se áreas importantes para a classificação. Desta forma, dever-se-á ter removido dados de treino que caracterizavam subclasses ou zonas com mistura de classe não permitindo a correta classificação destas áreas. Isto parece sugerir que, quando se quer obter um mapa com algum nível de generalização, é mais adequado que os dados de treino conterem pixéis mistos.

Os MUCS obtidos a partir dos dados de treino provenientes da COS 2015 são melhores que os MUCS obtidos a partir dos dados de treino extraídos do OSM, quando comparados com um mapa de referência obtido a partir da COS 2018. Este facto era de esperar, pois a COS 2015 e a COS 2018 são muito semelhantes, havendo pouca área onde ocorreu mudança entre as duas versões. Desta forma, o mapa de referência criado a partir da COS 2018 poderá não ser o melhor para validar os MUCS obtidos a partir de dados de treino criados a partir do OSM.

De forma geral, os MUCS criados com os dados de treino obtidos em simultâneo a partir da COS 2015 e do OSM são melhores que os obtidos com dados de treino criados exclusivamente pelo OSM, mas ao mesmo tempo são piores que os obtidos com dados de treino criados exclusivamente pela COS 2015.

Os MUCS criados com os dados de treino extraídos exclusivamente a partir do UA 2012 são melhores que os obtidos com dados de treino criados exclusivamente pelo OSM, mas ao mesmo tempo são piores que os obtidos com dados de treino criados exclusivamente pela COS 2015. Com estes dados, a escolha do critério para criação dos dados não tem grande influência na exatidão global dos mapas obtidos, havendo no máximo uma melhoria da exatidão global de 1,1%.

O uso adicional das *features* que caracterizam os pixéis vizinhos, usando DT_C1_F1-4 em vez de DT_C1_F1, mostrou ser benéfica de forma geral. Assim, é expectável que na classificação de imagens satélite, para classificar um pixel, seja importante o que o rodeia. Este facto encontra-se de acordo com a 1ª Lei de Tobler da Geografia: "Todas as coisas estão relacionadas com todas as outras, mas coisas próximas estão mais relacionadas do que coisas distantes" (Tobler, 1970).

Uma limitação associada aos produtos utilizados como fontes de dados de treino está relacionado com a sua disponibilidade. Desta forma, a COS 2015 tem a limitação de apenas existir para Portugal, enquanto que o OSM está disponível para todo o planeta. No entanto, apesar da sua extensão geográfica, tem regiões que têm poucos dados. Relativamente ao UA, este apenas existe para cidades com um mínimo de habitantes, não existindo dados para a área de estudo B que continha o Parque Natural da Serra da Estrela. Assim, tem-se como ordem de preferência de uso se existirem dados: COS 2015, UA 2012 e OSM.

Como trabalho futuro, dada a limitação de tempo para a realização deste trabalho e a limitação computacional, poder-se-á realizar as seguintes tarefas:

- Fazer uma análise pormenorizada dos erros de omissão e comissão para cada classe em cada MUCS produzido;
- Criar outras condições para o critério L de filtragem a partir dos índices radiométricos;

- Testar filtragens a partir dos índices radiométricos em dados de treino extraídos de outros produtos que não a COS 2015;
- Avaliar a separabilidade das classes duas a duas;
- Repetir todas as classificações feitas utilizando a totalidade dos dados de treino em vez de amostras utilizando o classificador *Random Forest* ou outros classificadores supervisionados como os baseados em redes neurais;
- Aumentar o número de vizinhos como *features*;
- Analisar as potencialidades do OSM usando também a aplicação dos índices radiométricos para filtragem ou como *features*;
- Validar os MUCS obtidos utilizando uma base de dados de referência com informação que não tenha sido alvo de generalização, de forma a conseguir avaliar a qualidade real das classificações obtidas ao nível do pixel.

Bibliografia

- Ahiablame, L., Sinha, T., Paul, M., Ji, J.-H., & Rajib, A. (2017). Streamflow response to potential land use and climate changes in the James River watershed, Upper Midwest United States. *Journal of Hydrology: Regional Studies*, 14, 150–166. <https://doi.org/10.1016/j.ejrh.2017.11.004>
- Albarakat, R., & Lakshmi, V. (2019). Comparison of Normalized Difference Vegetation Index Derived from Landsat, MODIS, and AVHRR for the Mesopotamian Marshes Between 2002 and 2018. *Remote Sensing*, 11(10), 1245. <https://doi.org/10.3390/rs11101245>
- Bédard, Y., Wright, E., & Rivest, S. (2015, Junho 16). *A Guide to Geospatial Data Quality*. <https://doi.org/10.13140/RG.2.1.2819.1763>
- Bede, B. (2013). *Mathematics of Fuzzy Sets and Fuzzy Logic*. Springer-Verlag. <https://doi.org/10.1007/978-3-642-35221-8>
- Borghuis, A. M., Chang, K., & Lee, H. Y. (2007). Comparison between automated and manual mapping of typhoon-triggered landslides from SPOT-5 imagery. *International Journal of Remote Sensing*, 28(8), 1843–1856. <https://doi.org/10.1080/01431160600935638>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Calazans Campelo, C. E., Bertolotto, M., & Corcoran, P. (Eds.). (2017). *Volunteered Geographic Information and the Future of Geospatial Data*: IGI Global. <https://doi.org/10.4018/978-1-5225-2446-5>
- Chunara, R., Andrews, J. R., & Brownstein, J. S. (2012). Social and News Media Enable Estimation of Epidemiological Patterns Early in the 2010 Haitian Cholera Outbreak. *The American Journal of Tropical Medicine and Hygiene*, 86(1), 39–45. <https://doi.org/10.4269/ajtmh.2012.11-0597>
- Copernicus. (2016). *Mapping Guide v4.7 for a European Urban Atlas*.
- Cutler, A., Cutler, D. R., & Stevens, J. R. (2012). Random Forests. Em C. Zhang & Y. Ma (Eds.), *Ensemble Machine Learning* (pp. 157–175). Springer US. https://doi.org/10.1007/978-1-4419-9326-7_5
- Dalezios, N. R. (2014). *REMOTE SENSING POTENTIAL IN HYDROLOGICAL SIMULATION*. 13.

- Delwart, S. (2015). *ESA Standard Document. 1*, 64.
- Direção-Geral do Território. (2018). *Especificações técnicas da Carta de uso e ocupação do solo de Portugal Continental para 1995, 2007, 2010 e 2015*. 103.
- Direção-Geral do Território. (2019). *Especificações técnicas da Carta de Uso e Ocupação do Solo (COS) de Portugal Continental para 2018*.
- Dlamini, S. N., Beloconi, A., Mabaso, S., Vounatsou, P., Impouma, B., & Fall, I. S. (2019). Review of remotely sensed data products for disease mapping and epidemiology. *Remote Sensing Applications: Society and Environment*, *14*, 108–118. <https://doi.org/10.1016/j.rsase.2019.02.005>
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, *120*, 25–36. <https://doi.org/10.1016/j.rse.2011.11.026>
- Fuchs, G., Andrienko, N., Andrienko, G., Bothe, S., & Stange, H. (2013). Tracing the German centennial flood in the stream of tweets: First lessons learned. *Proceedings of the Second ACM SIGSPATIAL International Workshop on Crowdsourced and Volunteered Geographic Information - GEOCROWD '13*, 31–38. <https://doi.org/10.1145/2534732.2534741>
- Gallaun, H., & Hirschmugl, M. (2017). *URBAN ATLAS 2012 VALIDATION REPORT*.
- Goodchild, M. F., & Li, L. (2012). Assuring the quality of volunteered geographic information. *Spatial Statistics*, *1*, 110–120. <https://doi.org/10.1016/j.spasta.2012.03.002>
- Guptill, S. C., & Morrison, J. L. (1995). *Elements of Spatial Data Quality*. Elsevier. <https://doi.org/10.1016/C2009-0-14900-0>
- Haklay, M. (2010). How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets. *Environment and Planning B: Planning and Design*, *37*(4), 682–703. <https://doi.org/10.1068/b35097>
- He, C., Shi, P., Xie, D., & Zhao, Y. (2010). Improving the normalized difference built-up index to map urban built-up areas using a semiautomatic segmentation approach. *Remote Sensing Letters*, *1*(4), 213–221. <https://doi.org/10.1080/01431161.2010.481681>
- Hersperger, A. M., Oliveira, E., Pagliarin, S., Palka, G., Verburg, P., Bolliger, J., & Grădinaru, S. (2018). Urban land-use change: The role of strategic spatial planning. *Global Environmental Change*, *51*, 32–42. <https://doi.org/10.1016/j.gloenvcha.2018.05.001>
- Htitiou, A., Boudhar, A., Lebrini, Y., Hadria, R., Lionboui, H., Elmansouri, L., Tychon, B., & Benabdelouahab, T. (2019). The Performance of Random Forest Classification Based on Phenological Metrics Derived from Sentinel-2 and Landsat 8 to Map Crop Cover in an

- Irrigated Semi-arid Region. *Remote Sensing in Earth Systems Sciences*, 2(4), 208–224. <https://doi.org/10.1007/s41976-019-00023-9>
- Hu, T., Xu, J.-F., Zhang, D.-R., Wang, J., & Zhang, Y.-Z. (2013). Hard and Soft Classification Method of Multi-Spectral Remote Sensing Image Based on Adaptive Thresholds. *Guang Pu Xue Yu Guang Pu Fen Xi*, 33, 1038–1042. [https://doi.org/10.3964/j.issn.1000-0593\(2013\)04-1038-05](https://doi.org/10.3964/j.issn.1000-0593(2013)04-1038-05)
 - Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25(3), 295–309. [https://doi.org/10.1016/0034-4257\(88\)90106-X](https://doi.org/10.1016/0034-4257(88)90106-X)
 - Ihlen, V. (2019). *Landsat 7 (L7) Data Users Handbook*. 151.
 - Karanam, D. H. K. (2018). *STUDY OF NORMALIZED DIFFERENCE BUILT-UP (NDBI) INDEX IN AUTOMATICALLY MAPPING URBAN AREAS FROM LANDSAT TM IMAGERY*. 7(2279), 9.
 - Kavzoglu, T., & Mather, P. M. (2020). *The Use of Feature Selection Techniques in the Context of Artificial Neural Networks*. 7.
 - Kidane, M., Bezie, A., Kesete, N., & Tolessa, T. (2019). The impact of land use and land cover (LULC) dynamics on soil erosion and sediment yield in Ethiopia. *Heliyon*, 5(12), e02981. <https://doi.org/10.1016/j.heliyon.2019.e02981>
 - Kuncheva, L. (2000). *Fuzzy Classifier Design*. Physica-Verlag Heidelberg. <https://doi.org/10.1007/978-3-7908-1850-5>
 - Lambin, E. F., Baulies, X., Bockstael, N., Fischer, G., Krug, T., Leemans, R., Moran, E. F., Rindfuss, R. R., Sato, Y., Skole, D., Li, B. L. T., & Vogel, C. (1999). *A core project of the International Geosphere-Biosphere Programme and the International Human Dimensions Programme on Global Environmental Change*. 126.
 - Li, Z., Liu, S., Tan, Z., Sohl, T. L., & Wu, Y. (2017). Simulating the effects of management practices on cropland soil organic carbon changes in the Temperate Prairies Ecoregion of the United States from 1980 to 2012. *Ecological Modelling*, 365, 68–79. <https://doi.org/10.1016/j.ecolmodel.2017.09.017>
 - Luo, N., Wan, T., Hao, H., & Lu, Q. (2019). Fusing High-Spatial-Resolution Remotely Sensed Imagery and OpenStreetMap Data for Land Cover Classification Over Urban Areas. *Remote Sensing*, 11(1), 88. <https://doi.org/10.3390/rs11010088>
 - MacEachren, A. M., Jaiswal, A., Robinson, A. C., Pezanowski, S., Saveliev, A., Mitra, P., Zhang, X., & Blanford, J. (2011). SensePlace2: GeoTwitter analytics support for situational awareness. *2011 IEEE Conference on Visual Analytics Science and Technology (VAST)*, 181–190. <https://doi.org/10.1109/VAST.2011.6102456>
 - McFeeters, S. (2013). Using the Normalized Difference Water Index (NDWI) within a Geographic Information System to Detect Swimming Pools for Mosquito Abatement: A

- Practical Approach. *Remote Sensing*, 5(7), 3544–3561. <https://doi.org/10.3390/rs5073544>
- Pabian, F., & Institute for Transuranium Elements. (2015). *Commercial satellite imagery as an evolving open-source verification technology: Emerging trends and their impact for nuclear nonproliferation analysis*. Publications Office. <http://dx.publications.europa.eu/10.2789/439135>
 - Patriarca, J., Fonte, C. C., Estima, J., de Almeida, J.-P., & Cardoso, A. (2019). Automatic conversion of OSM data into LULC maps: Comparing FOSS4G based approaches towards an enhanced performance. *Open Geospatial Data, Software and Standards*, 4(1), 11. <https://doi.org/10.1186/s40965-019-0070-2>
 - Pontius Jr, R. G. P., & Millones, M. (2011). Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing*, 32(15), 4407–4429. <https://doi.org/10.1080/01431161.2011.552923>
 - Rajib, A., & Merwade, V. (2017). Hydrologic response to future land use change in the Upper Mississippi River Basin by the end of 21st century. *Hydrological Processes*, 31(21), 3645–3661. <https://doi.org/10.1002/hyp.11282>
 - Rajib, M. A., Ahiablame, L., & Paul, M. (2016). Modeling the effects of future land use change on water quality under multiple scenarios: A case study of low-input agriculture with hay/pasture production. *Sustainability of Water Quality and Ecology*, 8, 50–66. <https://doi.org/10.1016/j.swaqe.2016.09.001>
 - Regmi, R. R., Saha, S. K., & Balla, M. K. (1995). *Geospatial Analysis of Land Use Land Cover Change Modeling at Phewa Lake Watershed of Nepal by Using Cellular Automata Markov Model*. 8.
 - Ren, W., Tian, H., Tao, B., Yang, J., Pan, S., Cai, W.-J., Lohrenz, S. E., He, R., & Hopkinson, C. S. (2015). Large increase in dissolved inorganic carbon flux from the Mississippi River to Gulf of Mexico due to climatic and anthropogenic changes over the 21st century. *Journal of Geophysical Research: Biogeosciences*, 120(4), 724–736. <https://doi.org/10.1002/2014JG002761>
 - Sakaki, T., Okazaki, M., & Matsuo, Y. (2010). *Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors*. 10.
 - Senaratne, H., Mobasher, A., Ali, A. L., Capineri, C., & Haklay, M. (Muki). (2017). A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, 31(1), 139–167. <https://doi.org/10.1080/13658816.2016.1189556>
 - Sharma, P., & Kaur, M. (2013). Classification in Pattern Recognition: A Review. *International Journal of Advanced Research in Computer Science and Software Engineering*.

- 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>
- Stehman, S. V., & Foody, G. M. (2019). Key issues in rigorous accuracy assessment of land cover products. *Remote Sensing of Environment*, 231, 111199. <https://doi.org/10.1016/j.rse.2019.05.018>
- Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, 46, 234–240. JSTOR. <https://doi.org/10.2307/143141>
- Zhang, Z., Zhang, X., Jiang, X., Xin, Q., Ao, Z., Zuo, Q., & Chen, L. (2019). Automated Surface Water Extraction Combining Sentinel-2 Imagery and OpenStreetMap Using Presence and Background Learning (PBL) Algorithm. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(10), 3784–3798. <https://doi.org/10.1109/JSTARS.2019.2936406>

Anexos

- **Anexo A** – Criação de Grelhas
- **Anexo B** – Preenchimento das Grelhas
- **Anexo C** – Criação das Áreas de Treino a Partir da COS 2015
- **Anexo D** – Criação das Amostras das Áreas de Treino a Partir da COS 2015
- **Anexo E** – MUCS de Referência
- **Anexo F** – Classificações e Respetivas Generalizações a Partir dos Dados de Treino Extraídos da COS 2015
- **Anexo G** – Histogramas dos Índices NDVI, NDWI e NDBI por Classe
- **Anexo H** – Classificações e Respetivas Generalizações a Partir dos Dados de Treino Obtidos pela Filtragem dos Dados de Treino Extraídos da COS 2015
- **Anexo I** – Classificações e Respetivas Generalizações a Partir dos Dados de Treino Extraídos do OSM
- **Anexo J** – Classificações e Respetivas Generalizações a Partir dos Dados de Treino Obtidos em Simultâneo da COS 2015 e do OSM
- **Anexo K** – Classificações e Respetivas Generalizações a Partir dos Dados de Treino Extraídos do UA 2012
- **Anexo L** – Classificações e Respetivas Generalizações a Partir dos Dados de Treino Extraídos da COS 2015 Utilizando os Vizinhos como *Features*

Anexo A – Criação de Grelhas

Neste anexo é apresentado o código em linguagem *Python* para a criação de grelhas vazias. Dada a grande dimensão geográfica de cada área de estudo, este processo foi dividido na criação de várias grelhas. Para cada uma das áreas de estudo o processo é semelhante, sendo criadas grelhas de dimensão 10x10m correspondentes a janelas de 500x500 pixéis, tornando assim o processo menos dispendioso a nível computacional. Inicialmente são criadas as grelhas vazias, sem qualquer atributo associado, recorrendo-se aos módulos de *osgeo* e *osr*.

Código Python 1:

```
1. from osgeo import gdal, ogr
2. import os
3. import glob
4. import numpy as np
5. import osr
6. from math import ceil
7.
8.
9. FOLDER          = '/Users/ismaeljesus/Desktop/MEIG_Tese/Tese_Pratica/'
10.
11. LOCS = ['lx_3/', 'se_3/']
12. for LOC in LOCS:
13.     FOLDER_DATA  = os.path.join(os.path.join(FOLDER, 'Dados/'), LOC)
14.     S2_FOLDERS   = sorted(glob.glob(os.path.join(FOLDER_DATA, 'S2**')))
15.     FOLDER_OUTPUT = os.path.join(os.path.join(FOLDER, 'Outputs/'), LOC)
16.     FOLDER_OUTPUT_CORTEBANDAS = os.path.join(FOLDER_OUTPUT, 'CorteBandas/')
17.     FOLDER_OUTPUT_GRIDS = os.path.join(FOLDER_OUTPUT, 'Grids/')
18.     if not os.path.exists(FOLDER_OUTPUT_GRIDS):
19.         os.makedirs(FOLDER_OUTPUT_GRIDS)
20.
21.     raster_ref=sorted(glob.glob(os.path.join(FOLDER_OUTPUT_CORTEBANDAS, '**.tif'
22. ))) [0]
23.     tamX=500
24.     tamY=500
25.     Image = gdal.Open(raster_ref, gdal.GA_ReadOnly)
26.     geoTransform = Image.GetGeoTransform()
27.     xmin = geoTransform[0]
28.     ymax = geoTransform[3]
29.     xmax = xmin + geoTransform[1] * Image.RasterXSize
30.     ymin = ymax + geoTransform[5] * Image.RasterYSize
31.
32.     targetprj = osr.SpatialReference(wkt = Image.GetProjection())
33.     rowsPrinc = Image.RasterYSize
34.     colsPrinc = Image.RasterXSize
35.
36.     nfishX=ceil(colsPrinc/tamX)
37.     nfishY=ceil(rowsPrinc/tamY)
38.
39.     lastFishX=colsPrinc - (nfishX-1)*tamX
40.     lastFishY=rowsPrinc - (nfishY-1)*tamY
41.
```

```

42.     colsByFishnet=[]
43.     for i in range(nfishX):
44.         colsByFishnet.append(tamX)
45.     colsByFishnet[-1]=lastFishX
46.
47.     rowsByFishnet=[]
48.     for i in range(nfishY):
49.         rowsByFishnet.append(tamY)
50.     rowsByFishnet[-1]=lastFishY
51.
52.     for j in range(nfishY):
53.         for i in range(nfishX):
54.             cols=colsByFishnet[i]
55.             rows=rowsByFishnet[j]
56.             colsAntes=sum(colsByFishnet[0:i])
57.             rowsAntes=sum(rowsByFishnet[0:j])
58.
59.             outputGridfn= os.path.join(FOLDER_OUTPUT_GRIDS, 'grid_'+str(i*nfishY
+j)+'.shp')
60.
61.             ringXleftOrigin = xmin + colsAntes*geoTransform[1]
62.             ringXrightOrigin = ringXleftOrigin + geoTransform[1]
63.             ringYtopOrigin = ymax - rowsAntes*abs(geoTransform[5])
64.             ringYbottomOrigin = ringYtopOrigin-abs(geoTransform[5])
65.
66.             outDriver = ogr.GetDriverByName('ESRI Shapefile')
67.             if os.path.exists(outputGridfn):
68.                 os.remove(outputGridfn)
69.             outDataSource = outDriver.CreateDataSource(outputGridfn)
70.
71.             outLayer = outDataSource.CreateLayer(outputGridfn, targetprj, geom_t
ype=ogr.wkbPolygon )
72.             featureDefn = outLayer.GetLayerDefn()
73.
74.             countcols = 0
75.             while countcols < cols:
76.                 countcols += 1
77.
78.                 ringYtop = ringYtopOrigin
79.                 ringYbottom =ringYbottomOrigin
80.                 countrows = 0
81.
82.                 while countrows < rows:
83.                     countrows += 1
84.                     ring = ogr.Geometry(ogr.wkbLinearRing)
85.                     ring.AddPoint(ringXleftOrigin, ringYtop)
86.                     ring.AddPoint(ringXrightOrigin, ringYtop)
87.                     ring.AddPoint(ringXrightOrigin, ringYbottom)
88.                     ring.AddPoint(ringXleftOrigin, ringYbottom)
89.                     ring.AddPoint(ringXleftOrigin, ringYtop)
90.                     poly = ogr.Geometry(ogr.wkbPolygon)
91.                     poly.AddGeometry(ring)
92.
93.                     outFeature = ogr.Feature(featureDefn)
94.                     outFeature.SetGeometry(poly)
95.                     outLayer.CreateFeature(outFeature)
96.                     outFeature = None
97.
98.                     ringYtop = ringYtop - abs(geoTransform[5])
99.                     ringYbottom = ringYbottom - abs(geoTransform[5])
100.
101.                     ringXleftOrigin = ringXleftOrigin + geoTransform[1]
102.                     ringXrightOrigin = ringXrightOrigin + geoTransform[1]
103.
104.                 outDataSource = None

```

Anexo B – Preenchimento das Grelhas

A cada um dos polígonos das grelhas é associada a área ocupada por cada classe de cada um dos produtos. Para isso, é necessário começar por criar os atributos que guardarão a informação pretendida. De seguida, intersesta-se cada um dos polígonos da grelha com cada um dos polígonos existentes na COS 2015, COS 2018, OSM2LULC e UA 2012. Quando a interseção não for vazia, ao atributo correspondente à classe do produto a intersestar, na célula da grelha, atualiza-se a área ocupada pela respetiva classe/produto. De seguida apresenta-se a implementação em *Python* deste processo para o caso da COS 2015, sendo o processo análogo para os restantes produtos.

Código Python 2:

```
1. from osgeo import gdal, ogr
2. import os
3. import glob
4. import numpy as np
5. import osr
6. from math import ceil
7. import geopandas as gpd
8. import fiona
9.
10.
11. def OurClassByCOS2015(cell):
12.     cell=cell.split('.')
13.     if cell[0]=='1':
14.         if cell[1]=='1':
15.             return 1
16.         elif cell[1]=='2':
17.             return 1
18.         elif cell[1]=='3':
19.             return 1
20.         elif cell[1]=='4':
21.             if cell[2]=='1':
22.                 return 3
23.             elif cell[2]=='2':
24.                 if cell[3]=='01':
25.                     return 3
26.             else:
27.                 return 1
28.     elif cell[0]=='2':
29.         if cell[1]=='1':
30.             return 2
31.         elif cell[1]=='2':
32.             return 2
33.         elif cell[1]=='3':
34.             return 3
35.         elif cell[1]=='4':
36.             if cell[2]=='1':
37.                 return 2
38.             elif cell[2]=='2':
39.                 return 2
40.             elif cell[2]=='3':
```

```

41.         return 2 #3
42.     elif cell[2]=='4':
43.         return 4
44. elif cell[0]=='3':
45.     if cell[1]=='1':
46.         return 4
47.     elif cell[1]=='2':
48.         if cell[2]=='1':
49.             return 3
50.         elif cell[2]=='2':
51.             return 5
52.     elif cell[1]=='3':
53.         return 6
54. elif cell[0]=='4':
55.     return 7
56. elif cell[0]=='5':
57.     return 8
58.
59.
60. FOLDER      = '/Users/ismaeljesus/Desktop/MEIG_Tese/Tese_Pratica/'
61.
62. LOCS = ['lx_3/', 'se_3/']
63.
64. FIELDS=['cos2015_1',
65.         'cos2015_2',
66.         'cos2015_3',
67.         'cos2015_4',
68.         'cos2015_5',
69.         'cos2015_6',
70.         'cos2015_7',
71.         'cos2015_8',]
72.
73. for LOC in LOCS:
74.     FOLDER_DATA = os.path.join(os.path.join(FOLDER, 'Dados/'), LOC)
75.     S2_FOLDERS = sorted(glob.glob(os.path.join(FOLDER_DATA, 'S2**')))
76.     FOLDER_OUTPUT = os.path.join(os.path.join(FOLDER, 'Outputs/'), LOC)
77.     FOLDER_OUTPUT_GRIDS = os.path.join(FOLDER_OUTPUT, 'Grids/')
78.     FOLDER_OUTPUT_CUT=os.path.join(FOLDER_OUTPUT, 'COS2015/')
79.     if not os.path.exists(FOLDER_OUTPUT_CUT):
80.         os.makedirs(FOLDER_OUTPUT_CUT)
81.
82.     limite= glob.glob(os.path.join(FOLDER_DATA, 'lmt_estudo/**/*.shp'))[0]
83.     cos2015= os.path.join(FOLDER, 'Dados/COS2015_PT_32629/COS2015_32629.shp')
84.     cos2015out = os.path.join(FOLDER_OUTPUT_CUT, 'COS2015.shp')
85.
86.     g1 = gpd.GeoDataFrame.from_file(cos2015)
87.     g2 = gpd.GeoDataFrame.from_file(limite)
88.     result= gpd.overlay(g1, g2, how='intersection')
89.
90.     df = gpd.GeoDataFrame(result, crs={'init': 'epsg:32629'},)
91.     df.to_file(cos2015out)
92.     df.head()
93.
94.     driver =ogr.GetDriverByName('ESRI Shapefile')
95.     dataInt = driver.Open(cos2015out)
96.     layerInt = dataInt.GetLayer(0)
97.
98.     grids=sorted(glob.glob(os.path.join(FOLDER_OUTPUT_GRIDS, '**.shp')))
99.
100.     for grid in grids:
101.         print(grid)
102.         dataGrid = driver.Open(grid, 1)
103.
104.         layerGrid = dataGrid.GetLayer(0)
105.         layerDefinitionGrid = layerGrid.GetLayerDefn()
106.         fieldsshp=[]

```

```

107.         for i in range(layerDefinitionGrid.GetFieldCount()):
108.             if layerDefinitionGrid.GetFieldDefn(i).GetName() not in fiel
dsshp:
109.                 fieldsshp.append(layerDefinitionGrid.GetFieldDefn(i).Get
Name())
110.
111.             for field in FIELDS:
112.                 if field not in fieldsshp:
113.                     layerGrid.CreateField(ogr.FieldDefn(field, ogr.OFTReal))
114.
115.             extent = layerGrid.GetExtent()
116.
117.             ring = ogr.Geometry(ogr.wkbLinearRing)
118.             ring.AddPoint(extent[0],extent[2])
119.             ring.AddPoint(extent[1], extent[2])
120.             ring.AddPoint(extent[1], extent[3])
121.             ring.AddPoint(extent[0], extent[3])
122.             ring.AddPoint(extent[0],extent[2])
123.             polyGrid = ogr.Geometry(ogr.wkbPolygon)
124.             polyGrid.AddGeometry(ring)
125.
126.             featureExtGrid = ogr.Feature(layerDefinitionGrid)
127.             featureExtGrid.SetGeometry(polyGrid)
128.             geomGridExt=featureExtGrid.GetGeometryRef()
129.
130.             for feature in layerInt:
131.                 geomFeature=feature.GetGeometryRef()
132.
133.                 if geomFeature.Intersects(geomGridExt):
134.                     classe=OurClassByCOS2015(feature['COS2015_V1'])
135.
136.                     NameField="cos2015_"+str(classe)
137.                     for cell in layerGrid:
138.                         geomCell = cell.GetGeometryRef()
139.                         if geomFeature.Intersects(geomCell):
140.                             intersection = geomFeature.Intersection(geomCell
)
141.                             #print('intersectaCelula')
142.                             area = round(intersection.GetArea()*1000)/1000.0
143.
144.                             #print(area)
145.                             if cell[NameField]!=None:
146.                                 area= area+cell[NameField]
147.                                 area=round(area*1000)/1000.0
148.                                 if area>100.00:
149.                                     area=100.00
150.
151.                                 cell.SetField(NameField, area)
152.                                 layerGrid.SetFeature(cell)
153.                             layerGrid.ResetReading()
154.                             layerInt.ResetReading()
155.
156.             dataGrid=None
157.             dataInt=None
158.             layerGrid=None
159.             layerInt=None

```


Anexo C – Criação das Áreas de Treino a Partir da COS 2015

Neste anexo é apresentado o código em linguagem *Python* para a criação dos dados de treino extraídos diretamente e exclusivamente da COS 2015. Nos códigos Python 3, 4 e 5 têm-se as criações das áreas de treino com os critérios C0 (Presença), C1 (Predominância) e C2 (Exclusividade), respectivamente.

Código Python 3:

```
1. from osgeo import gdal, ogr
2. import os
3. import glob
4. import numpy as np
5. import osr
6. from math import ceil
7. import geopandas as gpd
8. import fiona
9.
10.
11. # PASTA RAIZ
12. FOLDER = '/Users/ismaeljesus/Desktop/MEIG_Tese/Tese_Pratica/'
13.
14. # PASTAS DE LOCALIZACOES
15. LOCS = ['lx_3/', 'se_3/']
16.
17. FIELDS=['cos2015_1',
18.         'cos2015_2',
19.         'cos2015_3',
20.         'cos2015_4',
21.         'cos2015_5',
22.         'cos2015_6',
23.         'cos2015_7',
24.         'cos2015_8',]
25.
26. for LOC in LOCS:
27.     FOLDER_DATA = os.path.join(os.path.join(FOLDER, 'Dados/'), LOC)
28.     S2_FOLDERS = sorted(glob.glob(os.path.join(FOLDER_DATA, 'S2**')))
29.     FOLDER_OUTPUT = os.path.join(os.path.join(FOLDER, 'Outputs/'), LOC)
30.     FOLDER_OUTPUT_GRIDS = os.path.join(FOLDER_OUTPUT, 'Grids/')
31.     FOLDER_OUTPUT_CORTEBANDAS = os.path.join(FOLDER_OUTPUT, 'CorteBandas/')
32.     FOLDER_OUTPUT_DT_C0shp=os.path.join(FOLDER_OUTPUT, 'DT_C0_shp/')
33.     if not os.path.exists(FOLDER_OUTPUT_DT_C0shp):
34.         os.makedirs(FOLDER_OUTPUT_DT_C0shp)
35.     FOLDER_OUTPUT_DT_C0=os.path.join(FOLDER_OUTPUT, 'DT_C0/')
36.     if not os.path.exists(FOLDER_OUTPUT_DT_C0):
37.         os.makedirs(FOLDER_OUTPUT_DT_C0)
38.
39.     grids=sorted(glob.glob(os.path.join(FOLDER_OUTPUT_GRIDS, '**.shp')))
40.
41.     source_schema = fiona.open(grids[0]).schema
42.     source_driver = fiona.open(grids[0]).driver
43.     source_crs = fiona.open(grids[0]).crs
```

```

44.
45.     for field in FIELDS:
46.         classe=field.split('_')[-1]
47.         with fiona.open(os.path.join(FOLDER_OUTPUT_DT_C0shp, 'classe'+str(classe
)+'.shp'),
48.                             'w',
49.                             driver=source_driver,
50.                             crs=source_crs,
51.                             schema=source_schema) as shpout:
52.             for grid in grids:
53.                 with fiona.open(grid) as source:
54.                     for feature in source:
55.                         if feature['properties'][field]!=None:
56.                             if(feature['properties'][field]>00.00 ):
57.                                 shpout.write(feature)
58.                             print(str(classe) + ': ' + grid)
59.
60.     raster_ref=glob.glob(os.path.join(FOLDER_OUTPUT_CORTEBANDAS, '**.tif'))[0]
61.
62.     shapes_classes=glob.glob(os.path.join(FOLDER_OUTPUT_DT_C0shp, '**.shp'))
63.
64.     for shape_classe in shapes_classes:
65.         raster_out_classe=os.path.join(FOLDER_OUTPUT_DT_C0,
66.                                         shape_classe.split('/')[0]-
1].split('.')[0]+'.tif')
67.         gdalformat = 'GTiff'
68.         datatype = gdal.GDT_Byte
69.         burnVal = 1
70.
71.         Image = gdal.Open(raster_ref, gdal.GA_ReadOnly)
72.         Shapefile = ogr.Open(shape_classe)
73.         Shapefile_layer = Shapefile.GetLayer()
74.
75.         Output = gdal.GetDriverByName(gdalformat).Create(raster_out_classe,
76.                                                         Image.RasterXSize,
77.                                                         Image.RasterYSize,
78.                                                         1,
79.                                                         datatype,
80.                                                         options=['COMPRESS=DEFL
ATE'])
81.         Output.SetProjection(Image.GetProjectionRef())
82.         Output.SetGeoTransform(Image.GetGeoTransform())
83.
84.         Band = Output.GetRasterBand(1)
85.         Band.SetNoDataValue(0)
86.         Band.Fill(0)
87.
88.         gdal.RasterizeLayer(Output,
89.                             [1],
90.                             Shapefile_layer,
91.                             burn_values=[burnVal])
92.
93.         Band = None
94.         Output = None
95.         Image = None
96.         Shapefile = None
97.         Shapefile_layer = None
98.
99.         print("Concluido: ", raster_out_classe)

```

Código Python 4:

```

1. from osgeo import gdal, ogr

```

```

2. import os
3. import glob
4. import numpy as np
5. import osr
6. from math import ceil
7. import geopandas as gpd
8. import fiona
9.
10. FOLDER      = '/Users/ismaeljesus/Desktop/MEIG_Tese/Tese_Pratica/'
11.
12. LOCS = ['lx_3/', 'se_3/']
13.
14. FIELDS=['cos2015_1',
15.         'cos2015_2',
16.         'cos2015_3',
17.         'cos2015_4',
18.         'cos2015_5',
19.         'cos2015_6',
20.         'cos2015_7',
21.         'cos2015_8',]
22.
23. for LOC in LOCS:
24.     FOLDER_DATA = os.path.join(os.path.join(FOLDER, 'Dados/'), LOC)
25.     S2_FOLDERS  = sorted(glob.glob(os.path.join(FOLDER_DATA, 'S2*')))
26.     FOLDER_OUTPUT = os.path.join(os.path.join(FOLDER, 'Outputs/'), LOC)
27.     FOLDER_OUTPUT_GRIDS = os.path.join(FOLDER_OUTPUT, 'Grids/')
28.     FOLDER_OUTPUT_CORTEBANDAS = os.path.join(FOLDER_OUTPUT, 'CorteBandas/')
29.     FOLDER_OUTPUT_DT_C1shp=os.path.join(FOLDER_OUTPUT, 'DT_C1_shp/')
30.     if not os.path.exists(FOLDER_OUTPUT_DT_C1shp):
31.         os.makedirs(FOLDER_OUTPUT_DT_C1shp)
32.     FOLDER_OUTPUT_DT_C1=os.path.join(FOLDER_OUTPUT, 'DT_C1/')
33.     if not os.path.exists(FOLDER_OUTPUT_DT_C1):
34.         os.makedirs(FOLDER_OUTPUT_DT_C1)
35.
36.     grids=sorted(glob.glob(os.path.join(FOLDER_OUTPUT_GRIDS, '*.shp')))
37.
38.     source_schema = fiona.open(grids[0]).schema
39.     source_driver = fiona.open(grids[0]).driver
40.     source_crs = fiona.open(grids[0]).crs
41.
42.     def calcAreas(feature):
43.         lclasses=[0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0]
44.         tudo0=True
45.         for k in range(8):
46.             if feature['properties'][FIELDS[k]]!=None:
47.                 lclasses[k]=lclasses[k]+feature['properties'][FIELDS[k]]
48.                 tudo0=False
49.
50.         return lclasses, tudo0
51.
52.     def getfieldmax(feature):
53.         [listaAreas,tudo0] = calcAreas(feature)
54.         indicemax = listaAreas.index(max(listaAreas))
55.         return FIELDS[indicemax]
56.
57.     for field in FIELDS:
58.         classe=field.split('_')[-1]
59.         with fiona.open(os.path.join(FOLDER_OUTPUT_DT_C1shp, 'classe'+str(classe)
60.                                     )+'.shp'),
61.                       'w',
62.                       driver=source_driver,
63.                       crs=source_crs,
64.                       schema=source_schema) as shpout:
65.             for grid in grids:
66.                 with fiona.open(grid) as source:
67.                     for feature in source:

```

```

67.             if feature['properties'][field]!=None:
68.                 if(feature['properties'][field]>00.00 ):
69.                     fieldmax=getfieldmax(feature)
70.                     if fieldmax==field:
71.                         shpout.write(feature)
72.                 print(str(classe) + ': ' + grid)
73.
74. raster_ref=glob.glob(os.path.join(FOLDER_OUTPUT_CORTEBANDAS, '**.tif'))[0]
75. shapes_classes=glob.glob(os.path.join(FOLDER_OUTPUT_DT_C1shp, '**.shp'))
76.
77. for shape_classe in shapes_classes:
78.     raster_out_classe=os.path.join(FOLDER_OUTPUT_DT_C1,
79.                                     shape_classe.split('/')[1].split('.')[0]+''.tif')
80.
81.     gdalformat = 'GTiff'
82.     datatype = gdal.GDT_Byte
83.     burnVal = 1
84.
85.     Image = gdal.Open(raster_ref, gdal.GA_ReadOnly)
86.
87.     Shapefile = ogr.Open(shape_classe)
88.     Shapefile_layer = Shapefile.GetLayer()
89.
90.     Output = gdal.GetDriverByName(gdalformat).Create(raster_out_classe,
91.                                                       Image.RasterXSize,
92.                                                       Image.RasterYSize,
93.                                                       1,
94.                                                       datatype,
95.                                                       options=['COMPRESS=DEFL
ATE'])
96.     Output.SetProjection(Image.GetProjectionRef())
97.     Output.SetGeoTransform(Image.GetGeoTransform())
98.
99.     Band = Output.GetRasterBand(1)
100.    Band.SetNoDataValue(0)
101.    Band.Fill(0)
102.
103.    gdal.RasterizeLayer(Output,
104.                        [1],
105.                        Shapefile_layer,
106.                        burn_values=[burnVal])
107.
108.    Band = None
109.    Output = None
110.    Image = None
111.    Shapefile = None
112.    Shapefile_layer = None
113.
114.    print("Concluido: ", raster_out_classe)

```

Código Python 5:

```

1. from osgeo import gdal, ogr
2. import os
3. import glob
4. import numpy as np
5. import osr
6. from math import ceil
7. import geopandas as gpd
8. import fiona
9.
10. FOLDER = '/Users/ismaeljesus/Desktop/MEIG_Tese/Tese_Pratica/'

```

```

11. LOCS = ['lx_3/', 'se_3/']
12.
13. FIELDS=['cos2015_1',
14.         'cos2015_2',
15.         'cos2015_3',
16.         'cos2015_4',
17.         'cos2015_5',
18.         'cos2015_6',
19.         'cos2015_7',
20.         'cos2015_8',]
21.
22. for LOC in LOCS:
23.     FOLDER_DATA = os.path.join(os.path.join(FOLDER, 'Dados/'), LOC)
24.     S2_FOLDERS = sorted(glob.glob(os.path.join(FOLDER_DATA, 'S2**')))
25.     FOLDER_OUTPUT = os.path.join(os.path.join(FOLDER, 'Outputs/'), LOC)
26.     FOLDER_OUTPUT_GRIDS = os.path.join(FOLDER_OUTPUT, 'Grids/')
27.     FOLDER_OUTPUT_CORTEBANDAS = os.path.join(FOLDER_OUTPUT, 'CorteBandas/')
28.     FOLDER_OUTPUT_DT_C2shp=os.path.join(FOLDER_OUTPUT, 'DT_C2_shp/')
29.     if not os.path.exists(FOLDER_OUTPUT_DT_C2shp):
30.         os.makedirs(FOLDER_OUTPUT_DT_C2shp)
31.     FOLDER_OUTPUT_DT_C2=os.path.join(FOLDER_OUTPUT, 'DT_C2/')
32.     if not os.path.exists(FOLDER_OUTPUT_DT_C2):
33.         os.makedirs(FOLDER_OUTPUT_DT_C2)
34.
35.     grids=sorted(glob.glob(os.path.join(FOLDER_OUTPUT_GRIDS, '**.shp')))
36.
37.     source_schema = fiona.open(grids[0]).schema
38.     source_driver = fiona.open(grids[0]).driver
39.     source_crs = fiona.open(grids[0]).crs
40.
41.     for field in FIELDS:
42.         classe=field.split('_')[-1]
43.         with fiona.open(os.path.join(FOLDER_OUTPUT_DT_C2shp, 'classe'+str(classe)
44.                                     )+'.shp'),
45.                        'w',
46.                        driver=source_driver,
47.                        crs=source_crs,
48.                        schema=source_schema) as shpout:
49.             for grid in grids:
50.                 with fiona.open(grid) as source:
51.                     for feature in source:
52.                         if feature['properties'][field]!=None:
53.                             if(feature['properties'][field]>=99.99):
54.                                 shpout.write(feature)
55.                                 print(str(classe) + ': ' + grid)
56.
57.     raster_ref=glob.glob(os.path.join(FOLDER_OUTPUT_CORTEBANDAS, '**.tif'))[0]
58.     shapes_classes=glob.glob(os.path.join(FOLDER_OUTPUT_DT_C2shp, '**.shp'))
59.     for shape_classe in shapes_classes:
60.         raster_out_classe=os.path.join(FOLDER_OUTPUT_DT_C2,
61.                                         shape_classe.split('/')[0]
62.                                         1].split('.')[0]+''.tif')
63.
64.         gdalformat = 'GTiff'
65.         datatype = gdal.GDT_Byte
66.         burnVal = 1
67.
68.         Image = gdal.Open(raster_ref, gdal.GA_ReadOnly)
69.         Shapefile = ogr.Open(shape_classe)
70.         Shapefile_layer = Shapefile.GetLayer()
71.
72.         Output = gdal.GetDriverByName(gdalformat).Create(raster_out_classe,
73.                                                         Image.RasterXSize,
74.                                                         Image.RasterYSize,
75.                                                         1,

```

```

75.                                     datatype,
76.                                     options=['COMPRESS=DEFL
    ATE'])
77.     Output.SetProjection(Image.GetProjectionRef())
78.     Output.SetGeoTransform(Image.GetGeoTransform())
79.
80.     Band = Output.GetRasterBand(1)
81.     Band.SetNoDataValue(0)
82.     Band.Fill(0)
83.
84.     gdal.RasterizeLayer(Output,
85.                         [1],
86.                         Shapefile_layer,
87.                         burn_values=[burnVal])
88.
89.     Band = None
90.     Output = None
91.     Image = None
92.     Shapefile = None
93.     Shapefile_layer = None
94.
95.     print("Concluido: ", raster_out_classe)

```

Anexo D – Criação das Amostras das Áreas de Treino a Partir da COS 2015

Neste anexo apresenta-se o código em linguagem *Python* para a criação das amostras dos dados de treino extraídos diretamente e exclusivamente da COS 2015.

Código Python 6:

```
1. from osgeo import gdal, ogr
2. import os
3. import glob
4. import numpy as np
5. import osr
6. from math import ceil
7. import geopandas as gpd
8. import fiona
9.
10. import random
11. random.seed(0)
12. FOLDER = '/Users/ismaeljesus/Desktop/MEIG_Tese/Tese_Pratica/'
13.
14. LOCS = ['lx_3/', 'se_3/']
15.
16. for LOC in LOCS:
17.     FOLDER_DATA = os.path.join(os.path.join(FOLDER, 'Dados/'), LOC)
18.     S2_FOLDERS = sorted(glob.glob(os.path.join(FOLDER_DATA, 'S2*')))
19.     FOLDER_OUTPUT = os.path.join(os.path.join(FOLDER, 'Outputs/'), LOC)
20.     FOLDER_OUTPUT_DT_C0=os.path.join(FOLDER_OUTPUT, 'DT_C0/')
21.     FOLDER_OUTPUT_DT_C1=os.path.join(FOLDER_OUTPUT, 'DT_C1/')
22.     FOLDER_OUTPUT_DT_C2=os.path.join(FOLDER_OUTPUT, 'DT_C2/')
23.     FOLDER_OUTPUT_sample0=os.path.join(FOLDER_OUTPUT, 'SampleProp_DT_C0/')
24.     if not os.path.exists(FOLDER_OUTPUT_sample0):
25.         os.makedirs(FOLDER_OUTPUT_sample0)
26.     FOLDER_OUTPUT_sample1=os.path.join(FOLDER_OUTPUT, 'SampleProp_DT_C1/')
27.     if not os.path.exists(FOLDER_OUTPUT_sample1):
28.         os.makedirs(FOLDER_OUTPUT_sample1)
29.     FOLDER_OUTPUT_sample2=os.path.join(FOLDER_OUTPUT, 'SampleProp_DT_C2/')
30.     if not os.path.exists(FOLDER_OUTPUT_sample2):
31.         os.makedirs(FOLDER_OUTPUT_sample2)
32.
33.     classes0=sorted(glob.glob(os.path.join(FOLDER_OUTPUT_DT_C0, '**.tif')))
34.     classes1=sorted(glob.glob(os.path.join(FOLDER_OUTPUT_DT_C1, '**.tif')))
35.     classes2=sorted(glob.glob(os.path.join(FOLDER_OUTPUT_DT_C2, '**.tif')))
36.
37.     file_link = gdal.Open(classes2[0])
38.     c = file_link.GetRasterBand(1).ReadAsArray().astype(np.int16)
39.
40.     exclusividade=c
41.     for file in classes2[1:]:
42.         file_link = gdal.Open(file)
43.
44.         c = file_link.GetRasterBand(1).ReadAsArray().astype(np.int16)
45.
46.         exclusividade=exclusividade+c
```



```

47.
48.     samplesC2=np.copy(exclusividade)
49.     samplesC2[samplesC2!=0]=0
50.
51.     for file in classes2:
52.
53.         outfile_name=os.path.join(FOLDER_OUTPUT_sample2,
54.                                   file.split('/')[-1])
55.         file_link = gdal.Open(file)
56.
57.         b1 = file_link.GetRasterBand(1).ReadAsArray().astype(np.int16)
58.
59.         x_pixels = b1.shape[1]
60.         y_pixels = b1.shape[0]
61.
62.         dtSize=sum(sum(b1))
63.         new=np.where(b1==1)
64.         ind1= new[0]
65.         ind2= new[1]
66.
67.         randomlist = sorted(random.sample(range(dtSize), int(dtSize/10)))
68.         ind1new=ind1[randomlist]
69.         ind2new=ind2[randomlist]
70.
71.         b1= np.where(b1==0, b1, 0)
72.
73.         for k in range(len(ind1new)):
74.             b1[ind1new[k]][ind2new[k]]=1
75.         samplesC2=samplesC2+b1
76.
77.         driver = gdal.GetDriverByName('GTiff')
78.
79.         newclasse_data = driver.Create(outfile_name,x_pixels, y_pixels, 1,gdal.G
DT_Int16)
80.
81.         newclasse_data.GetRasterBand(1).WriteArray(b1)
82.         geotrans=file_link.GetGeoTransform()
83.         proj=file_link.GetProjection()
84.         newclasse_data.SetGeoTransform(geotrans)
85.         newclasse_data.SetProjection(proj)
86.
87.         newclasse_data.FlushCache()
88.         newclasse_data=None
89.
90.     samplesC1=np.copy(samplesC2)
91.     for file in classes1:
92.         outfile_name=os.path.join(FOLDER_OUTPUT_sample1,
93.                                   file.split('/')[-1])
94.
95.         file_link = gdal.Open(file)
96.         b1 = file_link.GetRasterBand(1).ReadAsArray().astype(np.int16)
97.         bcopy = file_link.GetRasterBand(1).ReadAsArray().astype(np.int16)
98.
99.         x_pixels = b1.shape[1]
100.        y_pixels = b1.shape[0]
101.
102.        bex=b1-exclusividade
103.        bex=np.where(bex==1, bex, 0)
104.        dtSize=sum(sum(bex))
105.
106.        new=np.where(bex==1)
107.        ind1= new[0]
108.        ind2= new[1]
109.        randomlist = sorted(random.sample(range(dtSize), int(dtSize/10))
)
110.        ind1new=ind1[randomlist]

```

```

111.         ind2new=ind2[randomlist]
112.
113.         b1= np.where(b1==0, b1, 0)
114.
115.         for k in range(len(ind1new)):
116.             b1[ind1new[k]][ind2new[k]]=1
117.             samplesC1=samplesC1+b1
118.
119.             samplesC1[samplesC1>1]=1
120.
121.             b1=b1+np.where(bcopy==samplesC2, bcopy, 0)
122.
123.             driver = gdal.GetDriverByName('GTiff')
124.             newclasse_data = driver.Create(outfile_name,x_pixels, y_pixels,
1,gdal.GDT_Int16)
125.
126.             newclasse_data.GetRasterBand(1).WriteArray(b1)
127.
128.             geotrans=file_link.GetGeoTransform()
129.             proj=file_link.GetProjection()
130.             newclasse_data.SetGeoTransform(geotrans)
131.             newclasse_data.SetProjection(proj)
132.
133.             newclasse_data.FlushCache()
134.             newclasse_data=None
135.
136.
137.
138.         for file in classes0:
139.             outfile_name=os.path.join(FOLDER_OUTPUT_sample0,
140.                                     file.split('/')[ -1])
141.             file_link = gdal.Open(file)
142.
143.             bcopy = file_link.GetRasterBand(1).ReadAsArray().astype(np.int16
)
144.
145.             x_pixels = bcopy.shape[1]
146.             y_pixels = bcopy.shape[0]
147.
148.             b1=np.where(bcopy==samplesC1, bcopy, 0)
149.             dtSize=sum(sum(b1))
150.
151.             driver = gdal.GetDriverByName('GTiff')
152.
153.             newclasse_data = driver.Create(outfile_name,x_pixels, y_pixels,
1,gdal.GDT_Int16)
154.
155.             newclasse_data.GetRasterBand(1).WriteArray(b1)
156.
157.             geotrans=file_link.GetGeoTransform()
158.             proj=file_link.GetProjection()
159.             newclasse_data.SetGeoTransform(geotrans)
160.             newclasse_data.SetProjection(proj)
161.
162.             newclasse_data.FlushCache()
163.             newclasse_data=None
164.

```


Anexo E – MUCS de Referência

Nas figuras que seguem são apresentados os mapas de referência para cada uma das áreas de estudo.

No caso da área de estudo A, esta área apresenta dois MUCS de referência, uma vez que para os MUCS criados com dados de treino extraídos do UA 2012 a nomenclatura adotada é diferente. Assim, para os MUCS criados a partir dos dados de treino extraídos do UA 2012, utiliza-se o Mapa de Referência 2 para o processo de validação. Para os restantes utiliza-se o Mapa de Referência 1.

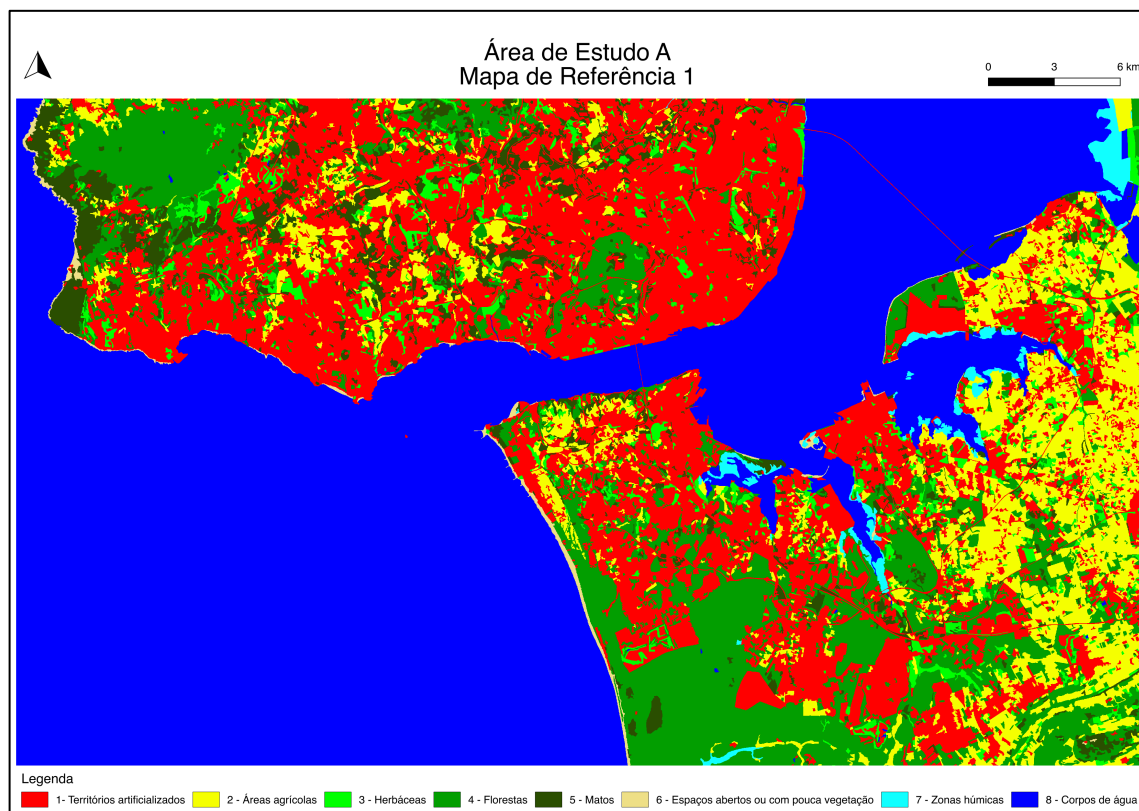


Figura 42: Área de Estudo A – Mapa de Referência 1.

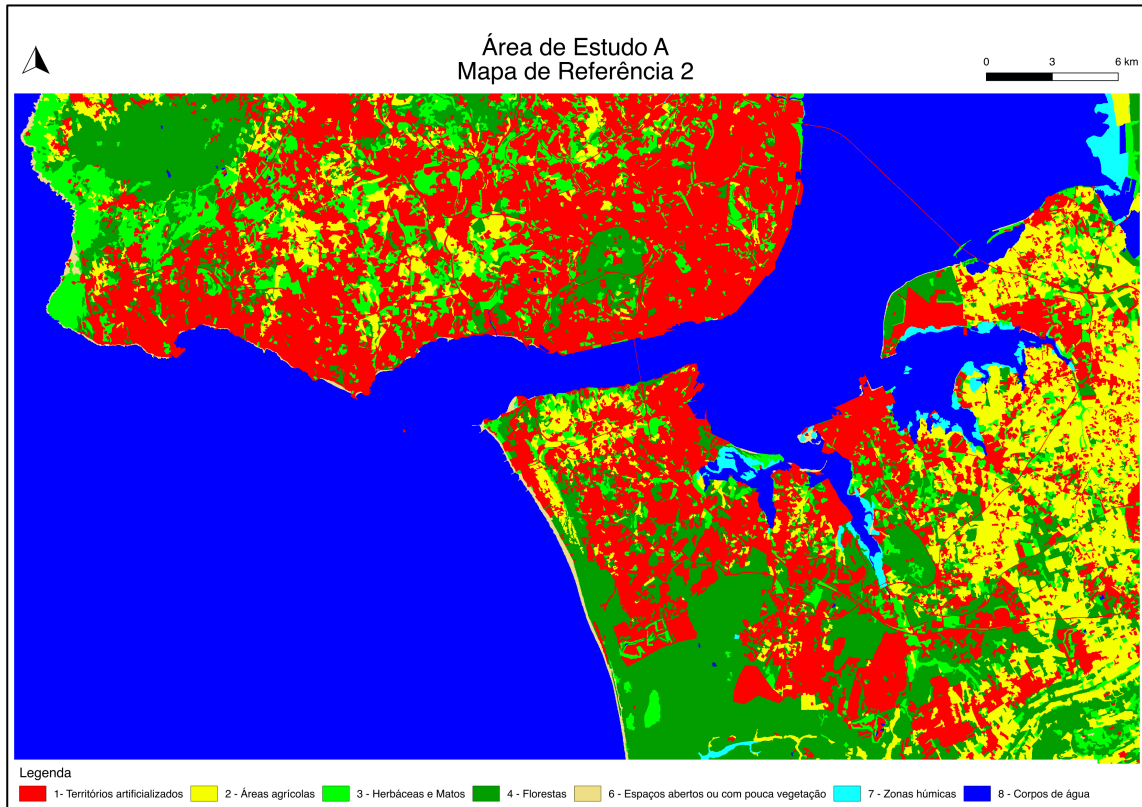


Figura 43: Área de Estudo A – Mapa de Referência 2.

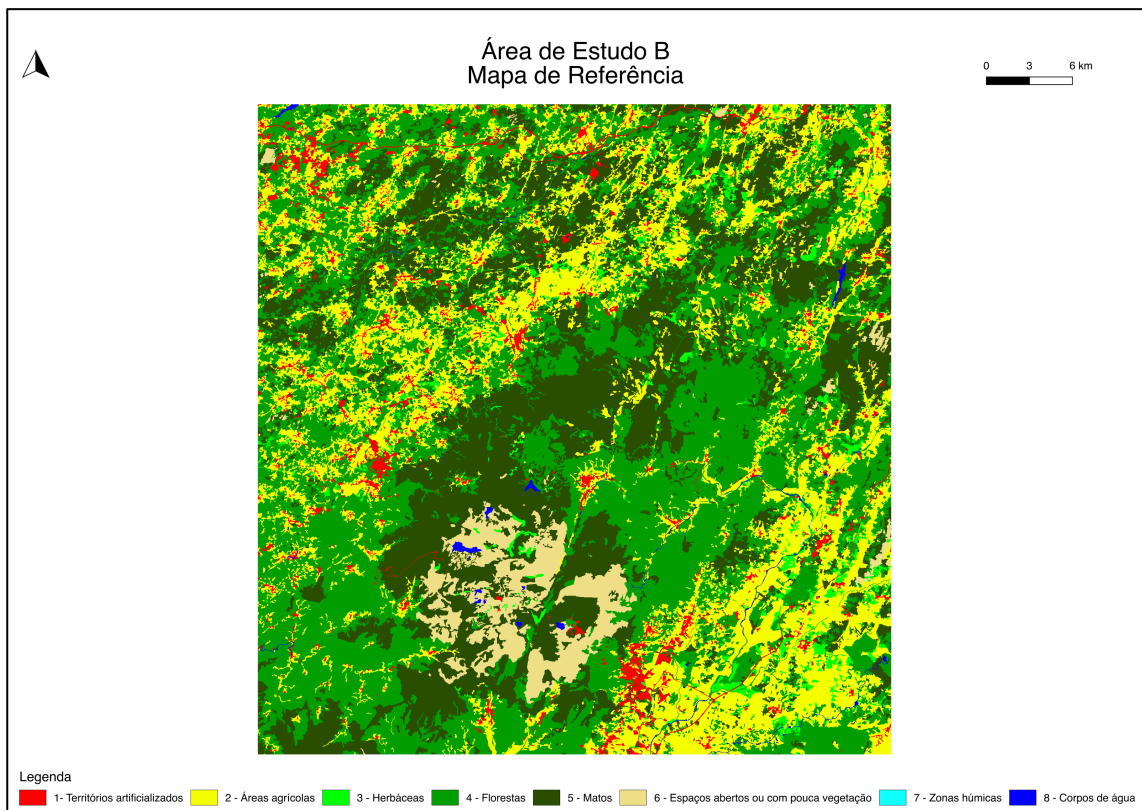


Figura 44: Área de Estudo B – Mapa de Referência.

Anexo F – Classificações e Respetivas Generalizações a Partir dos Dados de Treino Extraídos da COS 2015

Neste anexo apresentam-se os MUCS obtidos com os dados de treino extraídos da COS 2015 e as respetivas matrizes de confusão.

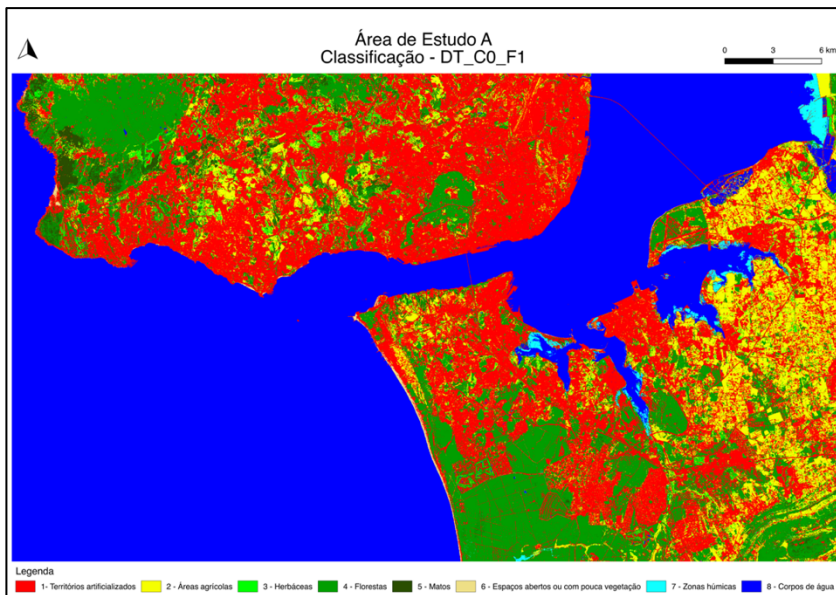


Figura 45: Área de Estudo A – MUCS obtido pela classificação com DT_C0_F1.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	21,5%	1,1%	0,2%	1,2%	0,1%	0,0%	0,0%	0,1%	89%
2	2,3%	5,5%	0,4%	0,7%	0,2%	0,0%	0,0%	0,0%	60%
3	0,9%	0,9%	0,8%	0,4%	0,1%	0,0%	0,0%	0,0%	25%
4	1,7%	0,5%	0,1%	11,1%	0,3%	0,0%	0,0%	0,0%	81%
5	1,0%	0,6%	0,3%	1,2%	1,2%	0,0%	0,0%	0,0%	29%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	38%
7	0,1%	0,0%	0,0%	0,1%	0,0%	0,0%	0,5%	0,1%	66%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,1%	43,9%	99%
Ex. Utilizador	78%	63%	41%	75%	62%	83%	85%	99%	85%

Tabela 55: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C0_F1.

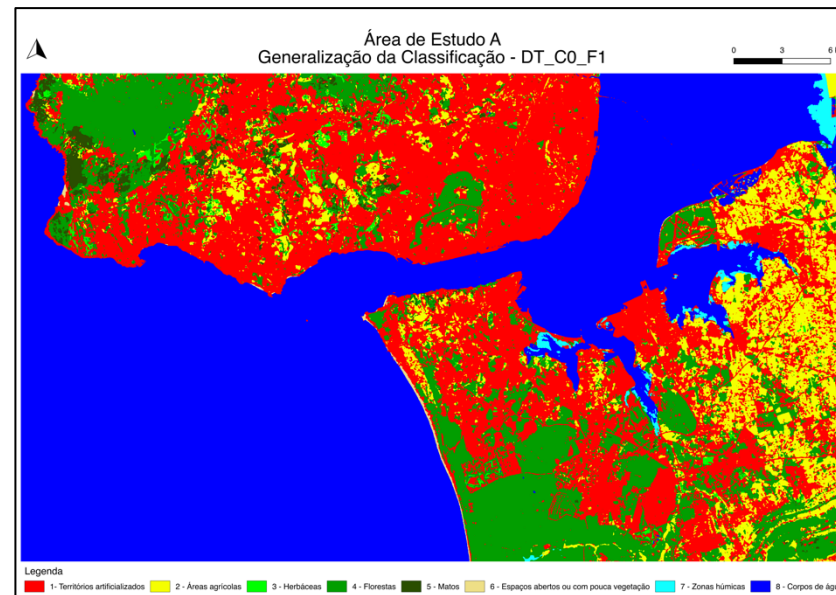


Figura 46: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C0_F1.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	23,0%	0,5%	0,0%	0,7%	0,0%	0,0%	0,0%	0,1%	95%
2	2,4%	6,0%	0,1%	0,6%	0,1%	0,0%	0,0%	0,0%	66%
3	1,2%	0,9%	0,6%	0,4%	0,0%	0,0%	0,0%	0,0%	19%
4	1,5%	0,3%	0,0%	11,8%	0,1%	0,0%	0,0%	0,0%	86%
5	1,4%	0,5%	0,1%	1,2%	1,1%	0,0%	0,0%	0,0%	25%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,1%	32%
7	0,1%	0,0%	0,0%	0,1%	0,0%	0,0%	0,6%	0,1%	70%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	44,0%	100%
Ex. Utilizador	77%	73%	67%	79%	84%	95%	92%	99%	87%

Tabela 56: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C0_F1.

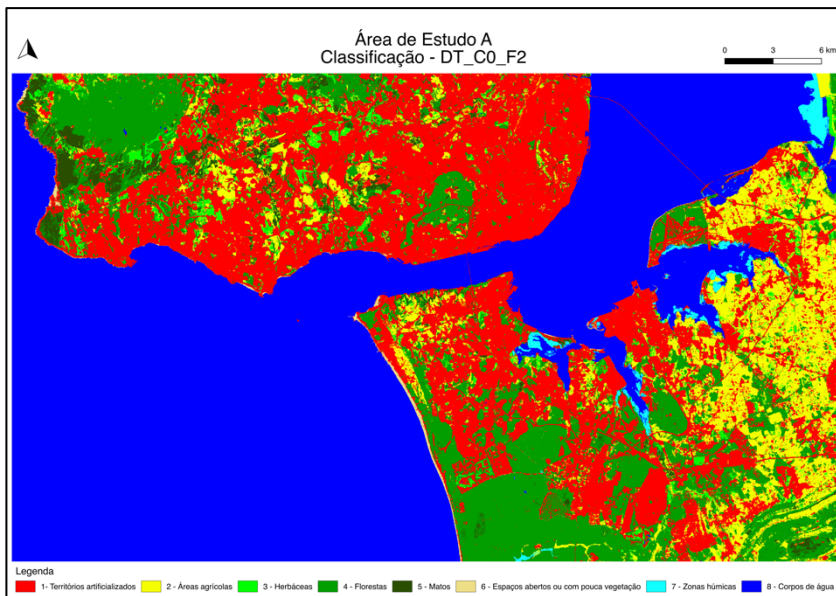


Figura 47: Área de Estudo A – MUCS obtido pela classificação com DT_C0_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	22,5%	0,7%	0,2%	0,8%	0,1%	0,0%	0,0%	0,0%	93%
2	1,2%	6,8%	0,5%	0,6%	0,1%	0,0%	0,0%	0,0%	74%
3	0,7%	0,7%	1,3%	0,3%	0,1%	0,0%	0,0%	0,0%	41%
4	1,0%	0,4%	0,1%	11,9%	0,2%	0,0%	0,0%	0,0%	87%
5	0,9%	0,4%	0,3%	0,8%	1,9%	0,0%	0,0%	0,0%	45%
6	0,1%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,0%	55%
7	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	80%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	44,0%	100%
Ex. Utilizador	85%	75%	55%	82%	78%	86%	92%	100%	89%

Tabela 57: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C0_F2.

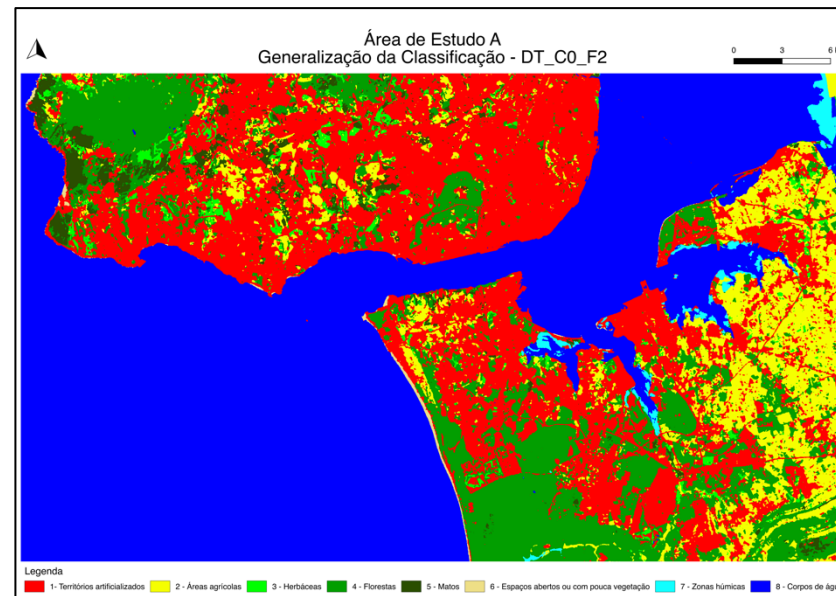


Figura 48: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C0_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	22,8%	0,5%	0,1%	0,7%	0,0%	0,0%	0,0%	0,1%	94%
2	1,3%	7,1%	0,3%	0,5%	0,1%	0,0%	0,0%	0,0%	77%
3	0,9%	0,7%	1,2%	0,3%	0,0%	0,0%	0,0%	0,0%	37%
4	1,1%	0,4%	0,1%	12,1%	0,1%	0,0%	0,0%	0,0%	88%
5	1,1%	0,3%	0,2%	0,8%	1,8%	0,0%	0,0%	0,0%	42%
6	0,1%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,0%	49%
7	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	81%
8	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	44,1%	100%
Ex. Utilizador	83%	78%	66%	83%	87%	93%	94%	100%	90%

Tabela 58: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C0_F2.

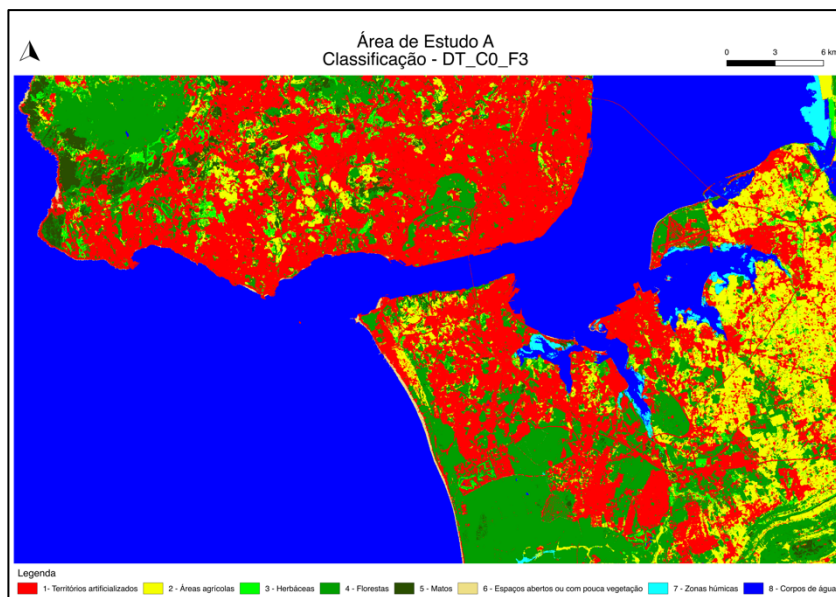


Figura 49: Área de Estudo A – MUCS obtido pela classificação com DT_C0_F3.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	22,4%	0,7%	0,2%	0,9%	0,1%	0,0%	0,0%	0,0%	92%
2	1,3%	6,7%	0,4%	0,6%	0,1%	0,0%	0,0%	0,0%	73%
3	0,8%	0,8%	1,2%	0,4%	0,1%	0,0%	0,0%	0,0%	37%
4	1,1%	0,5%	0,1%	11,8%	0,2%	0,0%	0,0%	0,0%	86%
5	0,9%	0,4%	0,3%	0,9%	1,8%	0,0%	0,0%	0,0%	42%
6	0,1%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,0%	52%
7	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	78%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	44,0%	100%
Ex. Utilizador	84%	74%	54%	81%	76%	86%	91%	100%	89%

Tabela 59: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C0_F3.

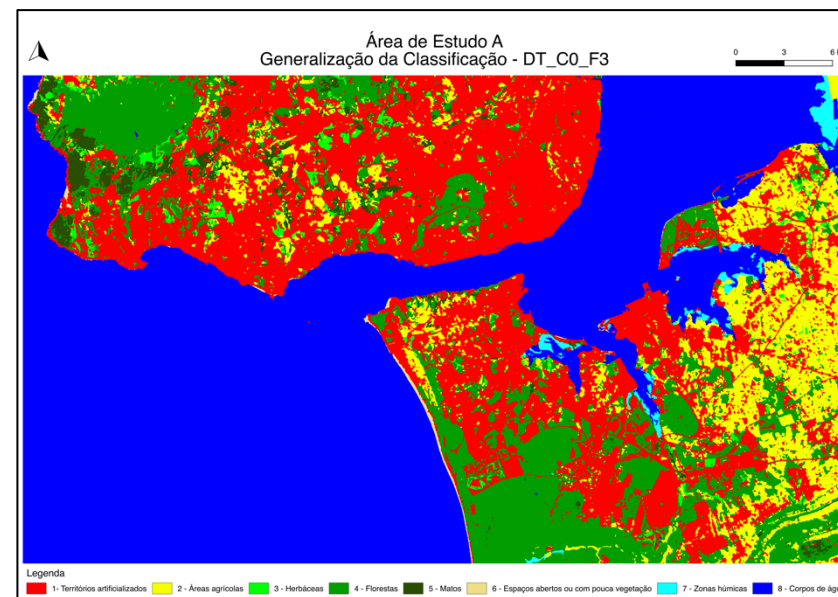


Figura 50: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C0_F3.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	22,8%	0,6%	0,1%	0,8%	0,0%	0,0%	0,0%	0,1%	94%
2	1,3%	7,0%	0,2%	0,6%	0,1%	0,0%	0,0%	0,0%	76%
3	0,9%	0,7%	1,1%	0,4%	0,0%	0,0%	0,0%	0,0%	33%
4	1,1%	0,4%	0,1%	12,0%	0,1%	0,0%	0,0%	0,0%	88%
5	1,2%	0,4%	0,2%	0,9%	1,7%	0,0%	0,0%	0,0%	39%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,1%	45%
7	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	79%
8	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	44,1%	100%
Ex. Utilizador	83%	77%	66%	82%	86%	93%	93%	99%	89%

Tabela 60: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C0_F3.

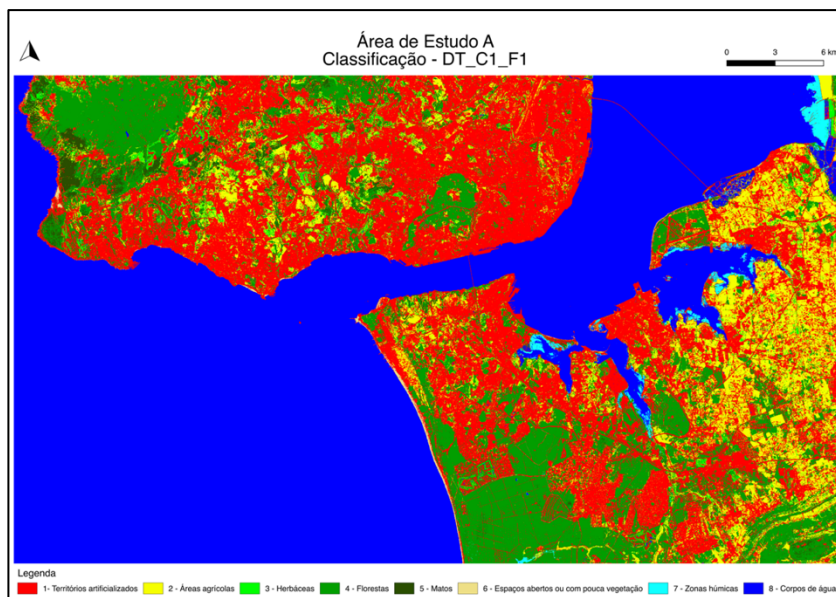


Figura 51: Área de Estudo A – MUCS obtido pela classificação com DT_C1_F1.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	21,5%	1,2%	0,2%	1,2%	0,1%	0,0%	0,0%	0,1%	88%
2	2,2%	5,6%	0,5%	0,7%	0,2%	0,0%	0,0%	0,0%	61%
3	0,9%	0,9%	0,8%	0,4%	0,1%	0,0%	0,0%	0,0%	26%
4	1,6%	0,6%	0,1%	11,1%	0,3%	0,0%	0,0%	0,0%	81%
5	0,9%	0,6%	0,3%	1,2%	1,3%	0,0%	0,0%	0,0%	29%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,0%	39%
7	0,1%	0,0%	0,0%	0,1%	0,0%	0,0%	0,5%	0,1%	67%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,1%	43,9%	99%
Ex. Utilizador	78%	63%	41%	76%	63%	84%	85%	99%	85%

Tabela 61: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1_F1.

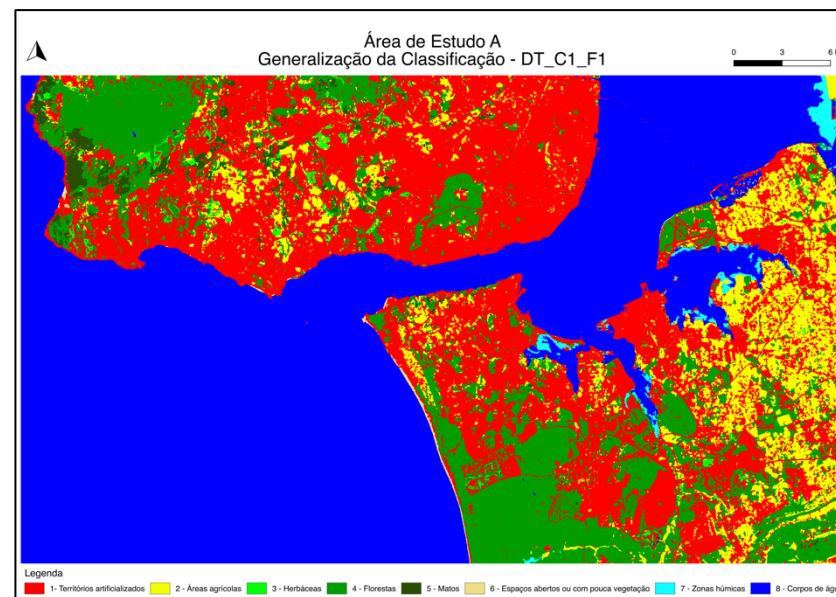


Figura 52: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C1_F1.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	23,0%	0,5%	0,0%	0,7%	0,0%	0,0%	0,0%	0,1%	95%
2	2,3%	6,1%	0,1%	0,6%	0,1%	0,0%	0,0%	0,0%	66%
3	1,1%	0,9%	0,6%	0,4%	0,0%	0,0%	0,0%	0,0%	19%
4	1,5%	0,4%	0,0%	11,8%	0,1%	0,0%	0,0%	0,0%	86%
5	1,3%	0,5%	0,1%	1,2%	1,1%	0,0%	0,0%	0,0%	25%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	33%
7	0,1%	0,0%	0,0%	0,1%	0,0%	0,0%	0,6%	0,1%	71%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	44,0%	100%
Ex. Utilizador	78%	72%	66%	79%	84%	95%	92%	99%	87%

Tabela 62: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1_F1.

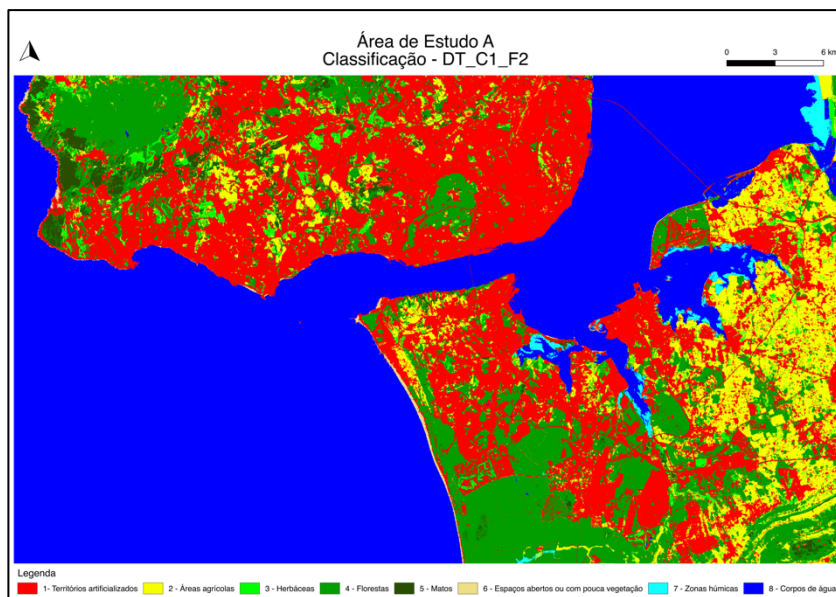


Figura 53: Área de Estudo A – MUCS obtido pela classificação com DT_C1_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	22,5%	0,7%	0,2%	0,8%	0,1%	0,0%	0,0%	0,0%	93%
2	1,2%	6,8%	0,5%	0,6%	0,1%	0,0%	0,0%	0,0%	74%
3	0,7%	0,7%	1,3%	0,3%	0,1%	0,0%	0,0%	0,0%	42%
4	1,0%	0,4%	0,1%	11,9%	0,2%	0,0%	0,0%	0,0%	87%
5	0,8%	0,4%	0,3%	0,8%	1,9%	0,0%	0,0%	0,0%	45%
6	0,1%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,0%	56%
7	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	80%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	44,0%	100%
Ex. Utilizador	85%	75%	55%	82%	78%	86%	92%	100%	89%

Tabela 63: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1_F2.

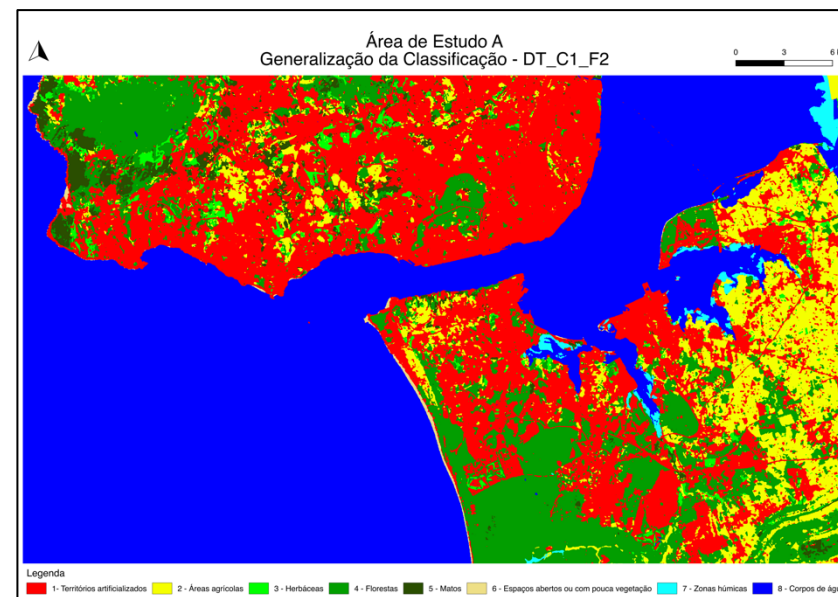


Figura 54: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C1_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	22,8%	0,5%	0,1%	0,7%	0,0%	0,0%	0,0%	0,1%	94%
2	1,2%	7,1%	0,3%	0,5%	0,1%	0,0%	0,0%	0,0%	77%
3	0,9%	0,7%	1,2%	0,3%	0,0%	0,0%	0,0%	0,0%	38%
4	1,1%	0,4%	0,1%	12,1%	0,1%	0,0%	0,0%	0,0%	88%
5	1,1%	0,4%	0,2%	0,8%	1,8%	0,0%	0,0%	0,0%	42%
6	0,1%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,0%	49%
7	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	82%
8	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	44,1%	100%
Ex. Utilizador	84%	78%	67%	83%	88%	93%	94%	100%	90%

Tabela 64: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1_F2.

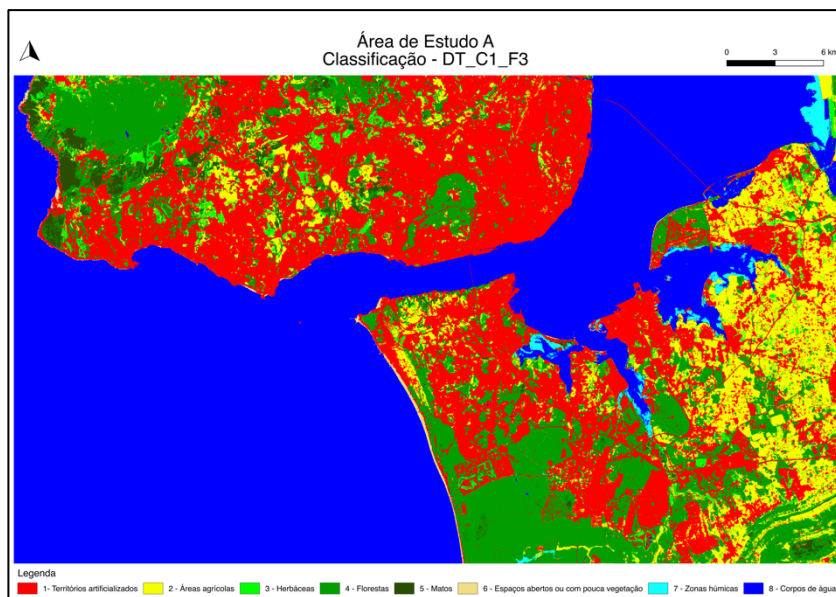


Figura 55: Área de Estudo A – MUCS obtido pela classificação com DT_C1_F3.

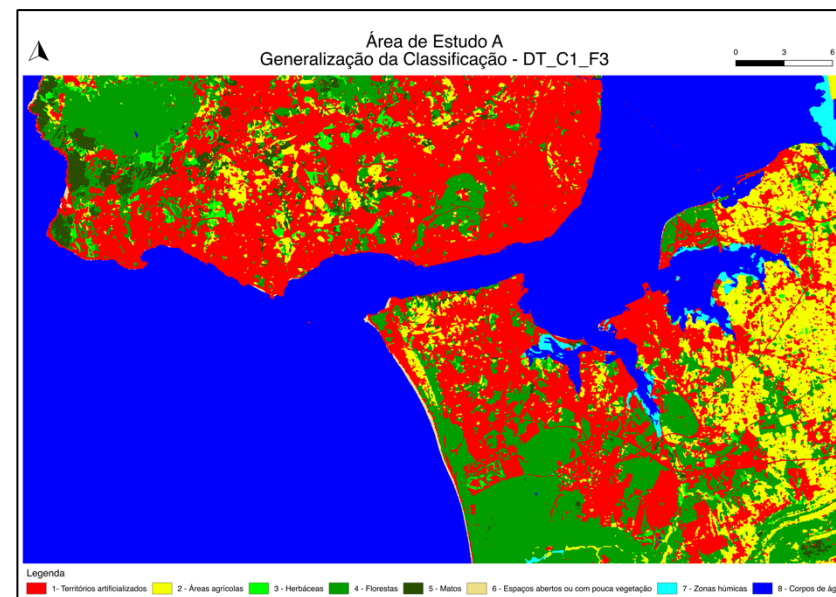


Figura 56: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C1_F3.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	22,4%	0,7%	0,2%	0,9%	0,1%	0,0%	0,0%	0,0%	92%
2	1,3%	6,7%	0,4%	0,6%	0,1%	0,0%	0,0%	0,0%	73%
3	0,7%	0,8%	1,2%	0,3%	0,1%	0,0%	0,0%	0,0%	38%
4	1,0%	0,5%	0,1%	11,8%	0,2%	0,0%	0,0%	0,0%	86%
5	0,8%	0,5%	0,3%	0,9%	1,8%	0,0%	0,0%	0,0%	42%
6	0,1%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,0%	53%
7	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	78%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	44,0%	100%
Ex. Utilizador	85%	73%	54%	81%	76%	87%	91%	100%	89%

Tabela 65: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1_F3.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	22,8%	0,6%	0,1%	0,8%	0,0%	0,0%	0,0%	0,1%	94%
2	1,3%	7,0%	0,2%	0,6%	0,1%	0,0%	0,0%	0,0%	77%
3	0,9%	0,8%	1,1%	0,4%	0,0%	0,0%	0,0%	0,0%	34%
4	1,1%	0,4%	0,1%	12,0%	0,1%	0,0%	0,0%	0,0%	88%
5	1,1%	0,4%	0,2%	0,9%	1,7%	0,0%	0,0%	0,0%	39%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,0%	46%
7	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	79%
8	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	44,1%	100%
Ex. Utilizador	83%	76%	66%	82%	86%	93%	93%	100%	89%

Tabela 66: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1_F3.

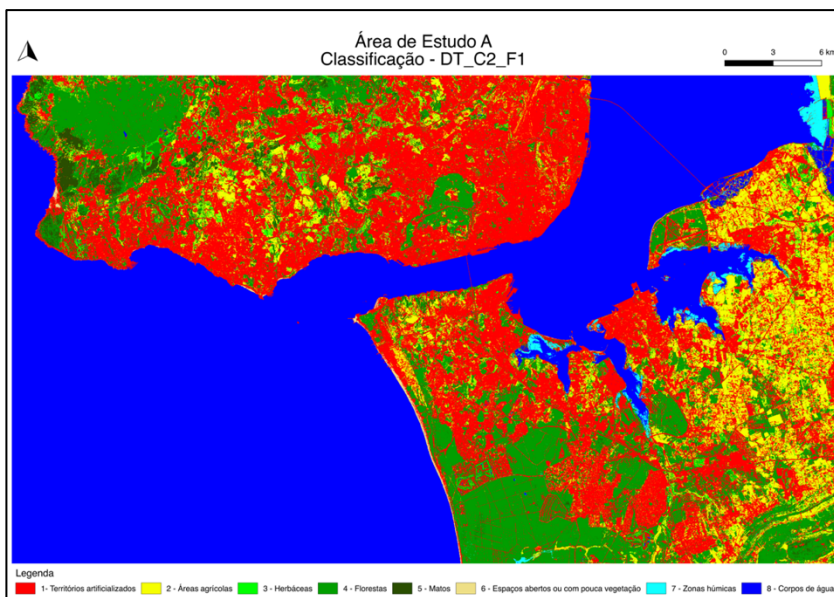


Figura 57: Área de Estudo A – MUCS obtido pela classificação com DT_C2_F1.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	21,5%	1,2%	0,2%	1,2%	0,1%	0,0%	0,0%	0,1%	88%
2	2,2%	5,6%	0,4%	0,7%	0,2%	0,0%	0,0%	0,0%	61%
3	0,9%	0,9%	0,8%	0,4%	0,1%	0,0%	0,0%	0,0%	25%
4	1,7%	0,6%	0,1%	11,0%	0,3%	0,0%	0,0%	0,0%	80%
5	0,9%	0,6%	0,3%	1,2%	1,2%	0,0%	0,0%	0,0%	29%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	37%
7	0,1%	0,0%	0,0%	0,1%	0,0%	0,0%	0,5%	0,1%	66%
8	0,2%	0,0%	0,0%	0,0%	0,0%	0,0%	0,1%	43,9%	99%
Ex. Utilizador	78%	62%	41%	76%	61%	85%	84%	99%	85%

Tabela 67: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C2_F1.

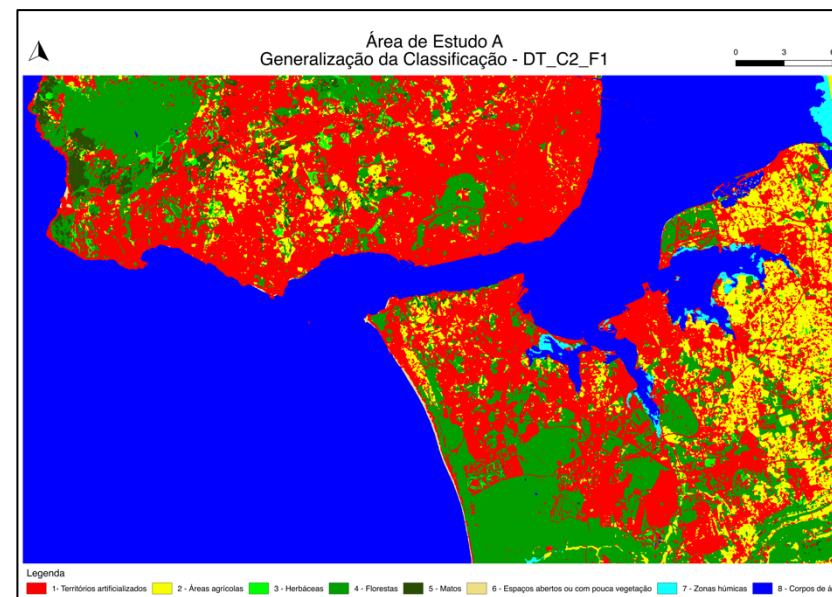


Figura 58: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C2_F1.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	23,0%	0,5%	0,0%	0,7%	0,0%	0,0%	0,0%	0,1%	95%
2	2,3%	6,1%	0,1%	0,6%	0,1%	0,0%	0,0%	0,0%	66%
3	1,1%	1,0%	0,6%	0,4%	0,0%	0,0%	0,0%	0,0%	19%
4	1,5%	0,4%	0,0%	11,7%	0,1%	0,0%	0,0%	0,0%	85%
5	1,3%	0,5%	0,1%	1,2%	1,1%	0,0%	0,0%	0,0%	25%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	31%
7	0,1%	0,0%	0,0%	0,1%	0,0%	0,0%	0,6%	0,1%	70%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	44,0%	100%
Ex. Utilizador	77%	72%	66%	79%	83%	95%	92%	99%	87%

Tabela 68: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C2_F1.

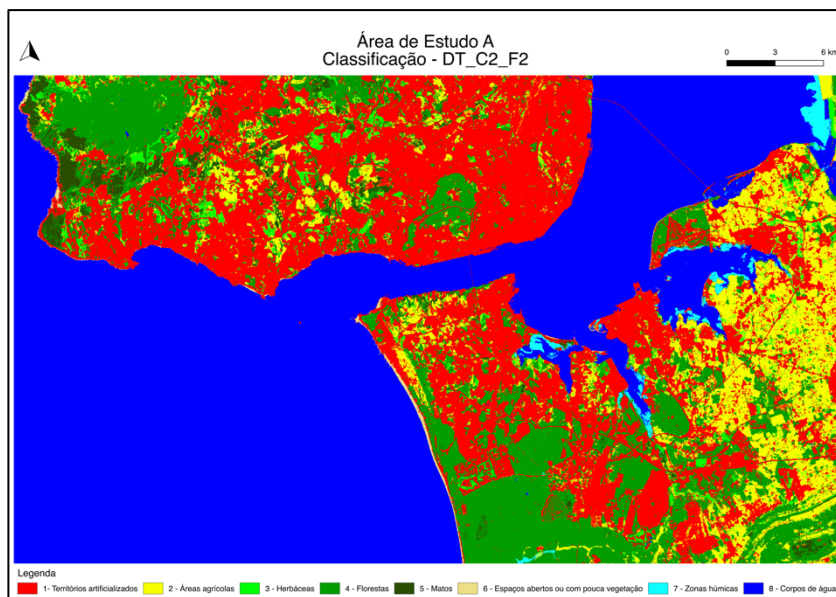


Figura 59: Área de Estudo A – MUCS obtido pela classificação com DT_C2_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	22,5%	0,7%	0,2%	0,8%	0,1%	0,0%	0,0%	0,0%	92%
2	1,2%	6,8%	0,4%	0,6%	0,1%	0,0%	0,0%	0,0%	74%
3	0,7%	0,7%	1,3%	0,3%	0,1%	0,0%	0,0%	0,0%	40%
4	1,0%	0,5%	0,1%	11,8%	0,2%	0,0%	0,0%	0,0%	86%
5	0,8%	0,4%	0,3%	0,8%	1,9%	0,0%	0,0%	0,0%	44%
6	0,1%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,0%	54%
7	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	79%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	44,0%	100%
Ex. Utilizador	85%	74%	55%	82%	77%	86%	91%	100%	89%

Tabela 69: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C2_F2.

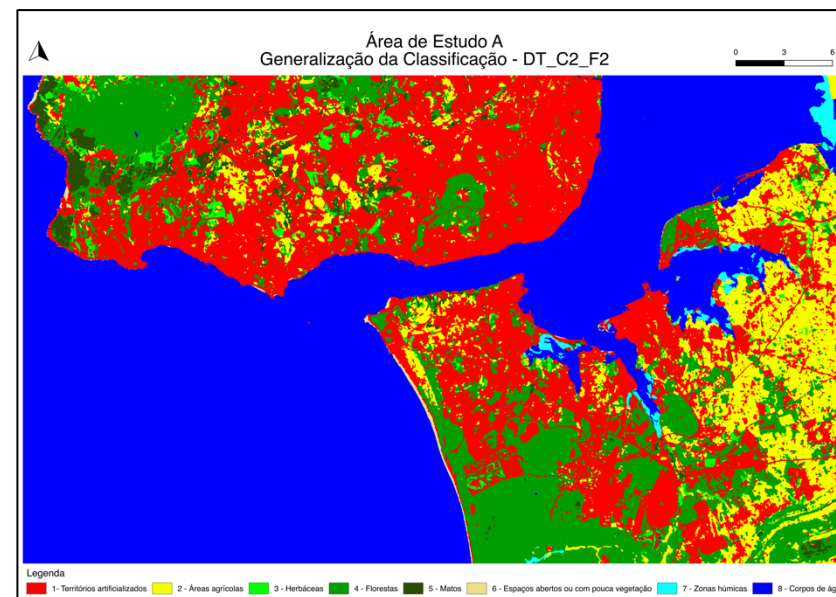


Figura 60: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C2_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	22,8%	0,6%	0,1%	0,8%	0,0%	0,0%	0,0%	0,1%	94%
2	1,2%	7,1%	0,2%	0,5%	0,1%	0,0%	0,0%	0,0%	77%
3	0,9%	0,7%	1,2%	0,3%	0,1%	0,0%	0,0%	0,0%	37%
4	1,1%	0,4%	0,1%	12,0%	0,1%	0,0%	0,0%	0,0%	88%
5	1,1%	0,4%	0,2%	0,8%	1,8%	0,0%	0,0%	0,0%	41%
6	0,1%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,0%	47%
7	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	81%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	44,1%	100%
Ex. Utilizador	83%	77%	66%	83%	87%	93%	94%	100%	90%

Tabela 70: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C2_F2.

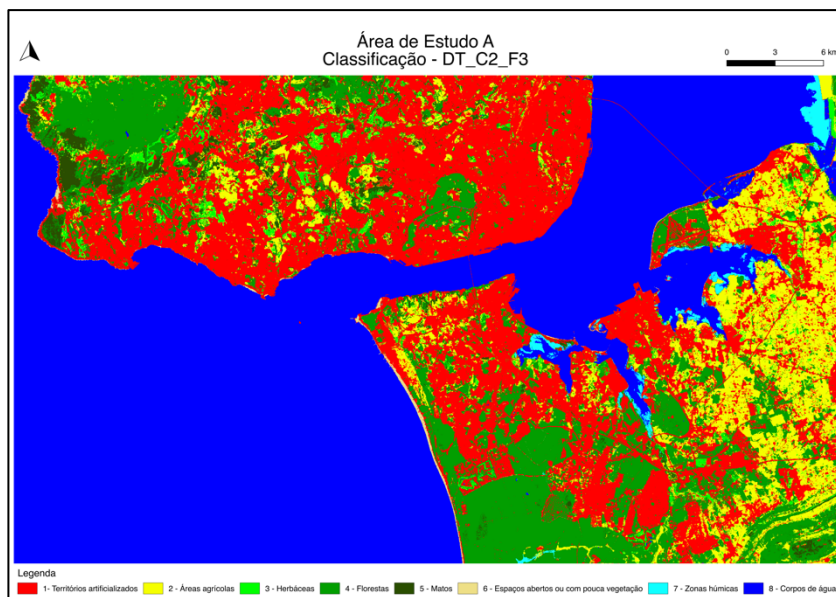


Figura 61: Área de Estudo A – MUCS obtido pela classificação com DT_C2_F3.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	22,4%	0,8%	0,2%	0,9%	0,1%	0,0%	0,0%	0,0%	92%
2	1,3%	6,7%	0,4%	0,6%	0,2%	0,0%	0,0%	0,0%	73%
3	0,7%	0,8%	1,1%	0,3%	0,1%	0,0%	0,0%	0,0%	36%
4	1,1%	0,5%	0,1%	11,8%	0,2%	0,0%	0,0%	0,0%	86%
5	0,9%	0,5%	0,3%	0,9%	1,8%	0,0%	0,0%	0,0%	41%
6	0,1%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,0%	51%
7	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	77%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,1%	44,0%	100%
Ex. Utilizador	84%	72%	53%	81%	75%	87%	90%	100%	89%

Tabela 71: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C2_F3.

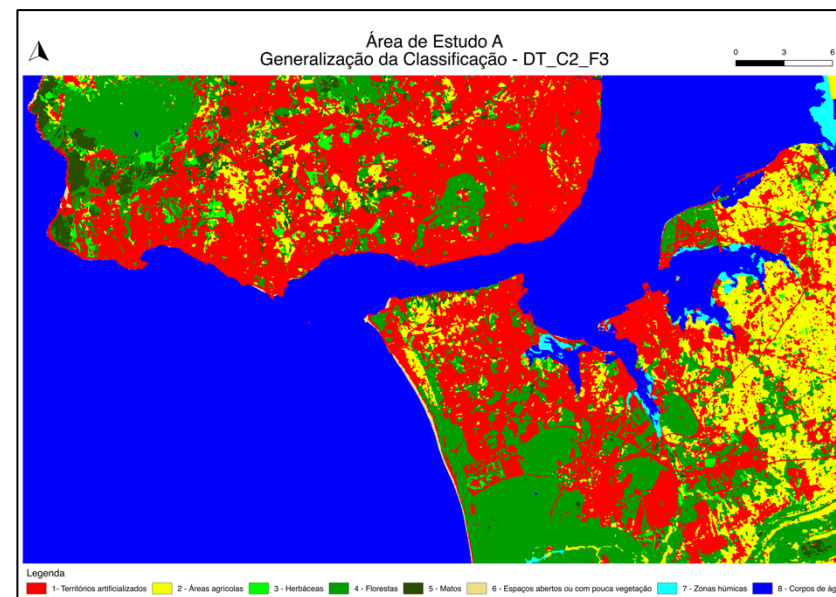


Figura 62: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C2_F3.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	22,7%	0,6%	0,1%	0,8%	0,0%	0,0%	0,0%	0,1%	93%
2	1,3%	7,0%	0,2%	0,6%	0,1%	0,0%	0,0%	0,0%	77%
3	0,9%	0,8%	1,0%	0,4%	0,1%	0,0%	0,0%	0,0%	33%
4	1,1%	0,4%	0,1%	12,0%	0,1%	0,0%	0,0%	0,0%	87%
5	1,1%	0,4%	0,2%	0,9%	1,7%	0,0%	0,0%	0,0%	39%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,0%	44%
7	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	79%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	44,1%	100%
Ex. Utilizador	83%	75%	66%	82%	86%	93%	93%	100%	89%

Tabela 72: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C2_F3.

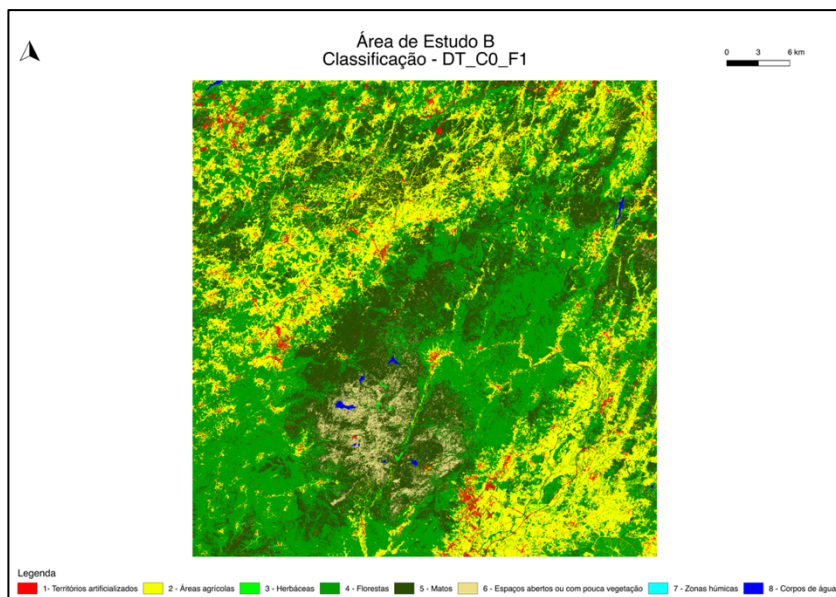


Figura 63: Área de Estudo B – MUCS obtido pela classificação com DT_C0_F1.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	1,5%	1,0%	0,0%	0,3%	0,4%	0,0%	0,0%	46%
2	0,3%	18,4%	0,2%	2,8%	1,4%	0,0%	0,0%	80%
3	0,0%	1,2%	0,2%	0,4%	0,6%	0,0%	0,0%	10%
4	0,1%	2,6%	0,0%	30,2%	5,8%	0,1%	0,0%	78%
5	0,1%	1,9%	0,1%	7,2%	18,3%	0,6%	0,0%	65%
6	0,0%	0,0%	0,0%	0,3%	2,0%	1,6%	0,0%	41%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	53%
Ex. Utilizador	72%	73%	44%	73%	65%	70%	90%	70%

Tabela 73: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C0_F1.

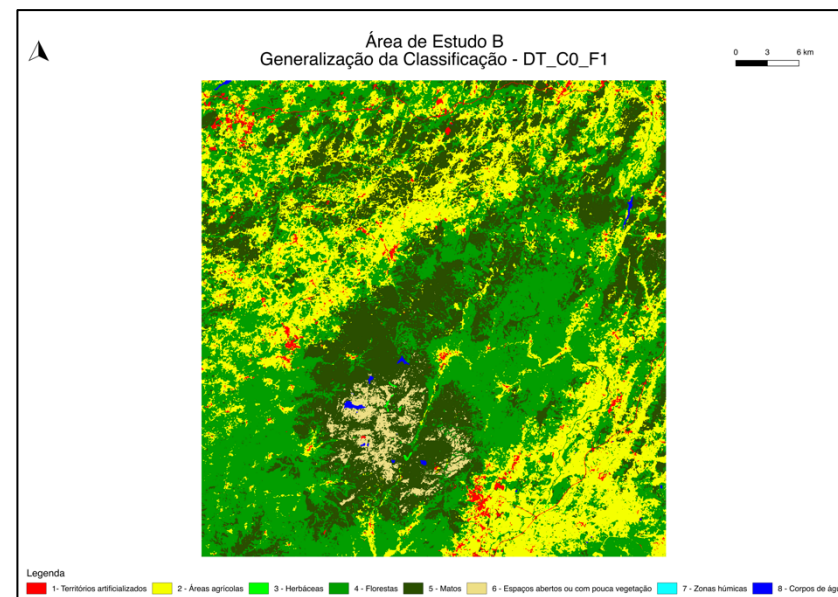


Figura 64: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C0_F1.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	1,5%	1,3%	0,0%	0,3%	0,1%	0,0%	0,0%	47%
2	0,1%	20,0%	0,0%	2,4%	0,6%	0,0%	0,0%	87%
3	0,0%	1,5%	0,1%	0,3%	0,5%	0,0%	0,0%	2%
4	0,1%	2,5%	0,0%	32,2%	4,0%	0,0%	0,0%	83%
5	0,0%	1,8%	0,0%	6,4%	19,6%	0,2%	0,0%	70%
6	0,0%	0,0%	0,0%	0,2%	2,2%	1,5%	0,0%	39%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,1%	50%
Ex. Utilizador	90%	73%	87%	77%	72%	86%	93%	75%

Tabela 74: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C0_F1.

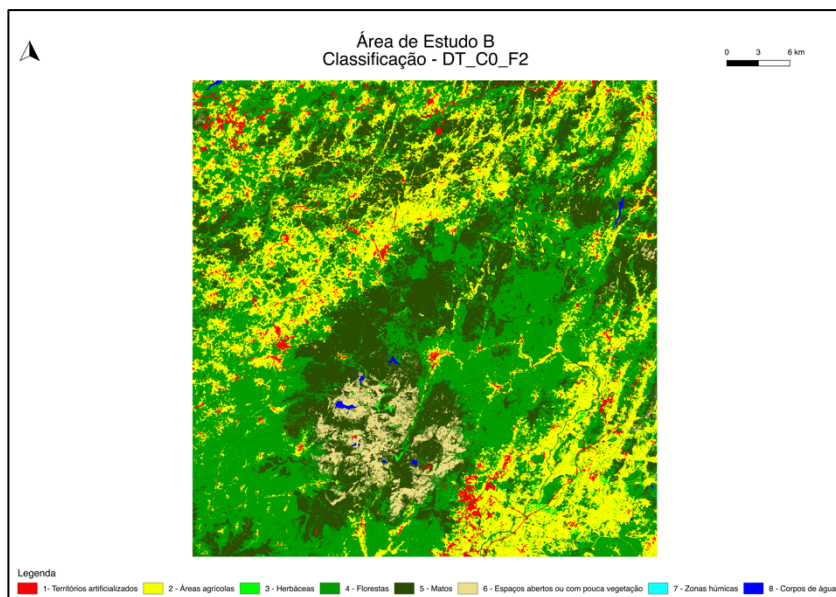


Figura 65: Área de Estudo B – MUCS obtido pela classificação com DT_C0_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,1%	0,8%	0,0%	0,2%	0,2%	0,0%	0,0%	64%
2	0,3%	19,4%	0,3%	2,2%	0,9%	0,0%	0,0%	84%
3	0,0%	1,1%	0,5%	0,3%	0,5%	0,0%	0,0%	21%
4	0,1%	2,1%	0,1%	32,6%	3,8%	0,0%	0,0%	84%
5	0,1%	1,3%	0,1%	4,9%	21,4%	0,3%	0,0%	76%
6	0,0%	0,0%	0,0%	0,2%	1,2%	2,6%	0,0%	66%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	58%
Ex. Utilizador	81%	78%	51%	80%	77%	87%	92%	79%

Tabela 75: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C0_F2.

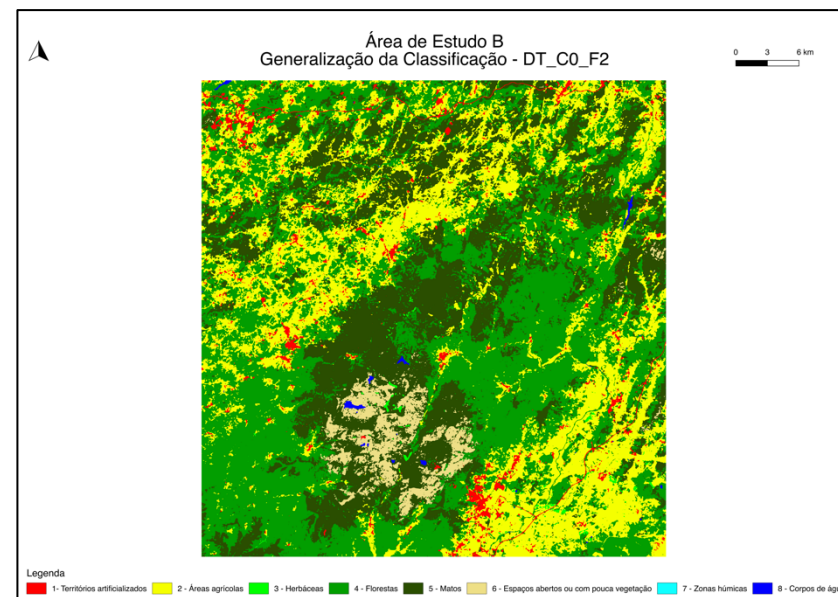


Figura 66: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C0_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,0%	0,9%	0,0%	0,2%	0,1%	0,0%	0,0%	63%
2	0,2%	20,0%	0,1%	2,2%	0,6%	0,0%	0,0%	87%
3	0,0%	1,3%	0,3%	0,3%	0,5%	0,0%	0,0%	12%
4	0,1%	2,2%	0,0%	33,5%	3,0%	0,0%	0,0%	86%
5	0,1%	1,3%	0,0%	4,3%	22,2%	0,2%	0,0%	79%
6	0,0%	0,0%	0,0%	0,2%	1,1%	2,7%	0,0%	68%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	52%
Ex. Utilizador	84%	78%	73%	82%	81%	91%	94%	81%

Tabela 76: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C0_F2.

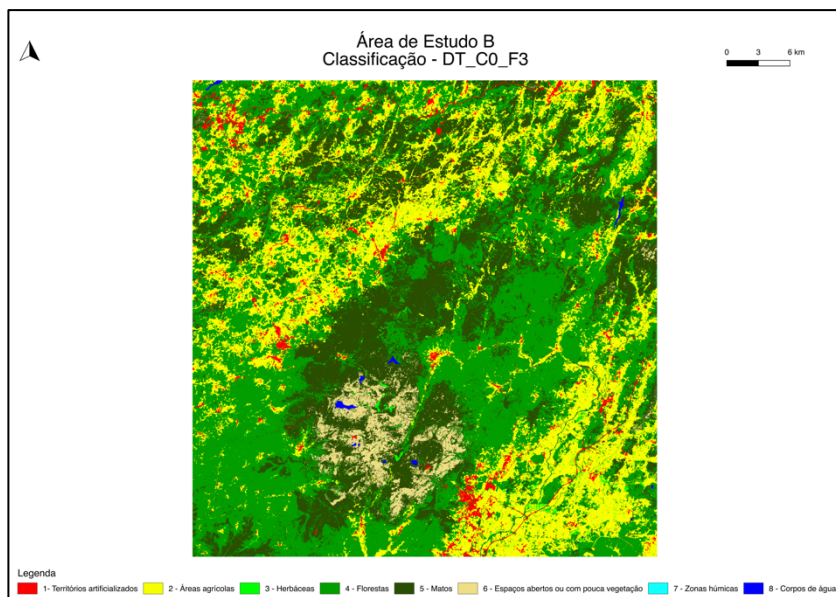


Figura 67: Área de Estudo B – MUCS obtido pela classificação com DT_C0_F3.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,0%	0,8%	0,0%	0,2%	0,2%	0,0%	0,0%	62%
2	0,3%	19,3%	0,2%	2,4%	0,9%	0,0%	0,0%	84%
3	0,0%	1,2%	0,4%	0,3%	0,5%	0,0%	0,0%	17%
4	0,1%	2,2%	0,0%	32,2%	4,2%	0,0%	0,0%	83%
5	0,1%	1,4%	0,1%	5,3%	20,9%	0,4%	0,0%	74%
6	0,0%	0,0%	0,0%	0,2%	1,3%	2,4%	0,0%	61%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	56%
Ex. Utilizador	80%	77%	52%	79%	75%	84%	91%	77%

Tabela 77: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C0_F3.

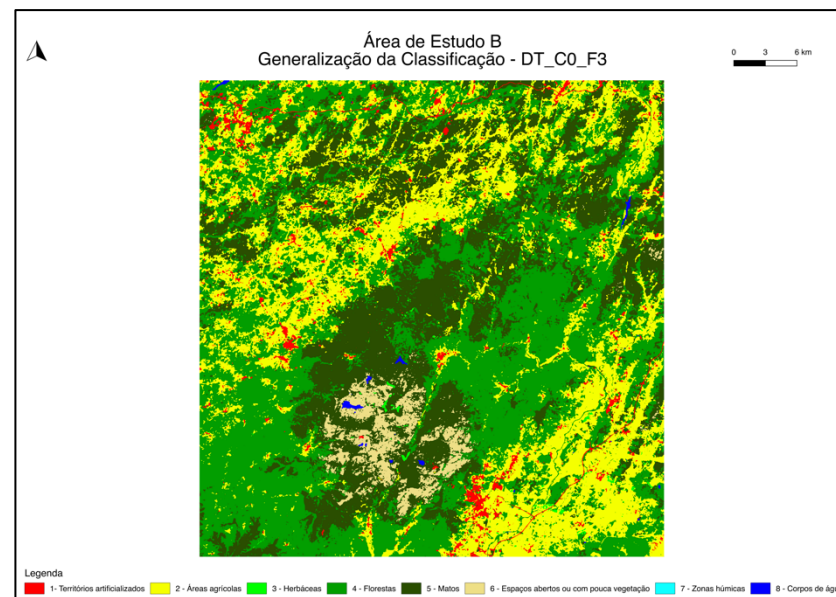


Figura 68: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C0_F3.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,0%	0,9%	0,0%	0,2%	0,1%	0,0%	0,0%	62%
2	0,2%	20,0%	0,0%	2,2%	0,6%	0,0%	0,0%	87%
3	0,0%	1,3%	0,3%	0,3%	0,5%	0,0%	0,0%	11%
4	0,1%	2,2%	0,0%	33,3%	3,1%	0,0%	0,0%	86%
5	0,1%	1,3%	0,0%	4,5%	22,0%	0,2%	0,0%	78%
6	0,0%	0,0%	0,0%	0,2%	1,2%	2,6%	0,0%	66%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	51%
Ex. Utilizador	84%	77%	75%	82%	80%	91%	94%	80%

Tabela 78: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C0_F3.

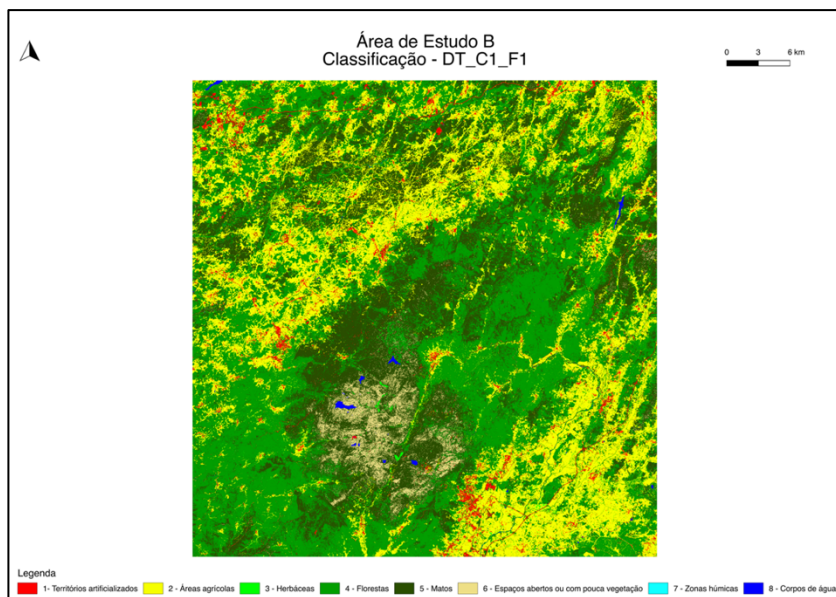


Figura 69: Área de Estudo B – MUCS obtido pela classificação com DT_C1_F1.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	1,6%	1,0%	0,0%	0,3%	0,4%	0,0%	0,0%	48%
2	0,3%	18,5%	0,2%	2,7%	1,3%	0,0%	0,0%	80%
3	0,0%	1,2%	0,3%	0,4%	0,6%	0,0%	0,0%	10%
4	0,1%	2,6%	0,0%	30,3%	5,7%	0,1%	0,0%	78%
5	0,1%	1,9%	0,1%	7,2%	18,3%	0,6%	0,0%	65%
6	0,0%	0,0%	0,0%	0,3%	2,0%	1,6%	0,0%	41%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	55%
Ex. Utilizador	72%	73%	45%	73%	65%	71%	91%	71%

Tabela 79: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1_F1.

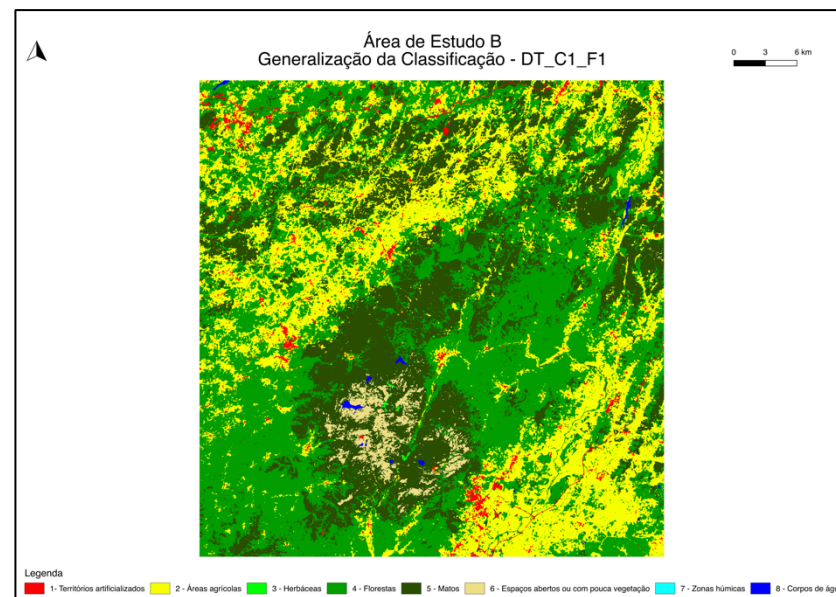


Figura 70: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C1_F1.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	1,5%	1,3%	0,0%	0,2%	0,1%	0,0%	0,0%	47%
2	0,1%	20,0%	0,0%	2,3%	0,6%	0,0%	0,0%	87%
3	0,0%	1,5%	0,1%	0,3%	0,6%	0,0%	0,0%	2%
4	0,1%	2,5%	0,0%	32,1%	4,1%	0,0%	0,0%	83%
5	0,0%	1,8%	0,0%	6,4%	19,7%	0,2%	0,0%	70%
6	0,0%	0,0%	0,0%	0,2%	2,2%	1,6%	0,0%	40%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,1%	50%
Ex. Utilizador	90%	73%	90%	77%	72%	86%	93%	75%

Tabela 80: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1_F1.

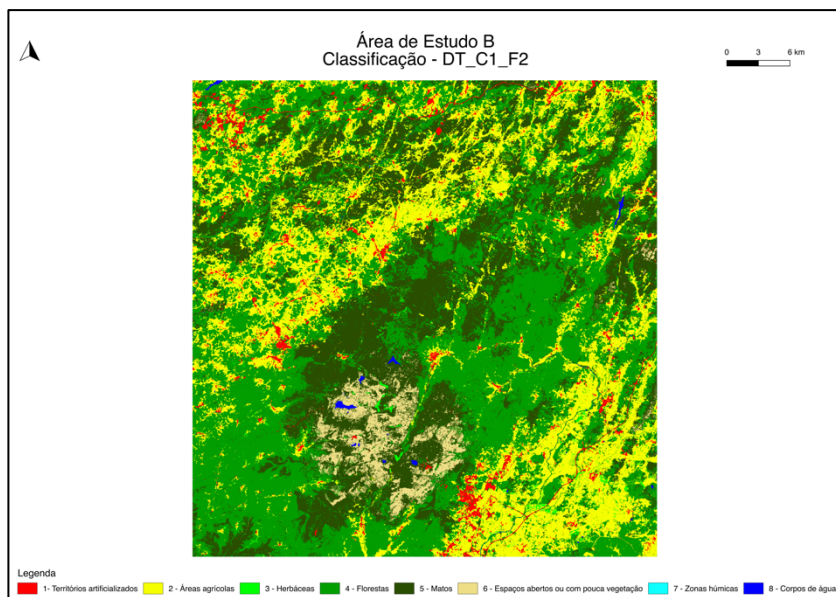


Figura 71: Área de Estudo B – MUCS obtido pela classificação com DT_C1_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,1%	0,7%	0,0%	0,2%	0,2%	0,0%	0,0%	65%
2	0,3%	19,4%	0,3%	2,2%	0,8%	0,0%	0,0%	84%
3	0,0%	1,1%	0,5%	0,3%	0,5%	0,0%	0,0%	22%
4	0,1%	2,1%	0,1%	32,7%	3,8%	0,0%	0,0%	84%
5	0,1%	1,3%	0,1%	5,0%	21,4%	0,3%	0,0%	76%
6	0,0%	0,0%	0,0%	0,2%	1,2%	2,6%	0,0%	66%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	59%
Ex. Utilizador	81%	79%	51%	81%	77%	87%	92%	79%

Tabela 81: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1_F2.

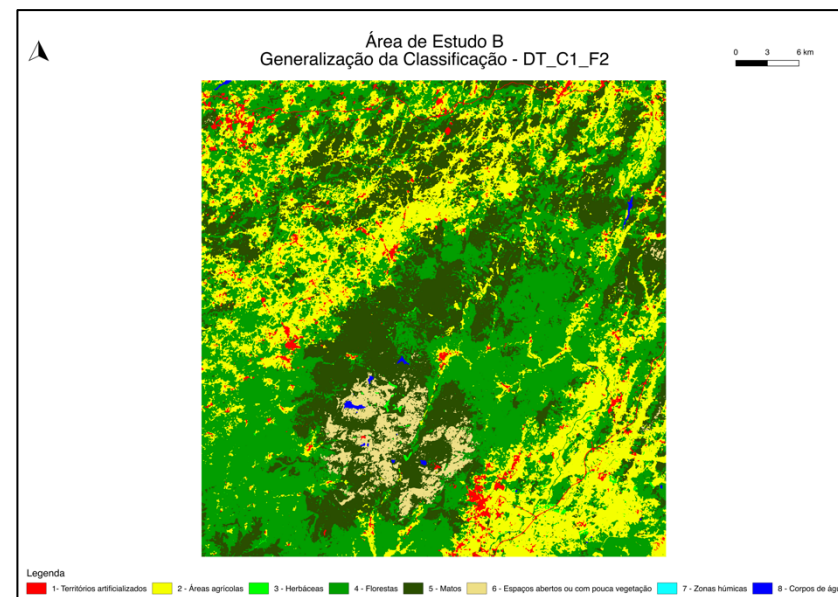


Figura 72: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C1_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,0%	0,9%	0,0%	0,2%	0,1%	0,0%	0,0%	63%
2	0,2%	20,0%	0,1%	2,2%	0,6%	0,0%	0,0%	87%
3	0,0%	1,3%	0,3%	0,3%	0,5%	0,0%	0,0%	12%
4	0,1%	2,2%	0,0%	33,3%	3,1%	0,0%	0,0%	86%
5	0,1%	1,3%	0,0%	4,4%	22,2%	0,2%	0,0%	79%
6	0,0%	0,0%	0,0%	0,1%	1,1%	2,7%	0,0%	67%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	52%
Ex. Utilizador	84%	78%	74%	82%	80%	91%	94%	81%

Tabela 82: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1_F2.

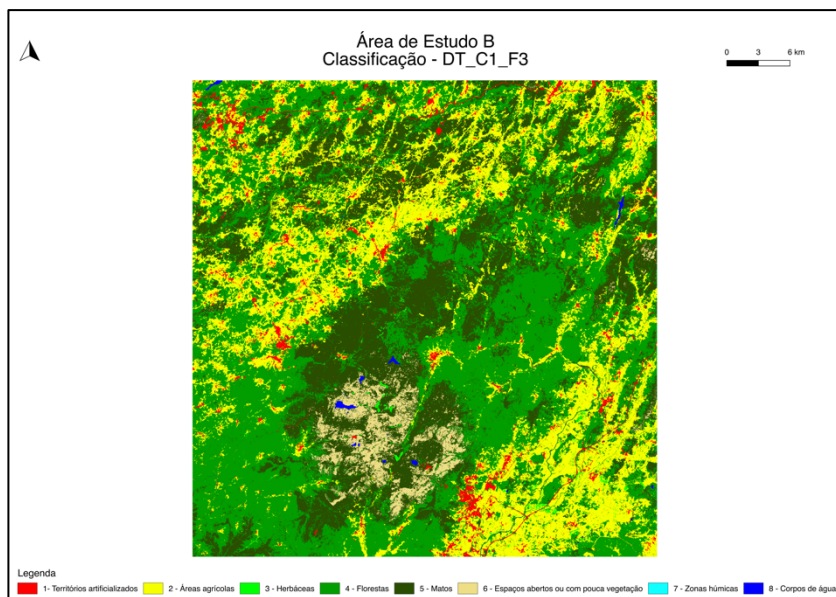


Figura 73: Área de Estudo B – MUCS obtido pela classificação com DT_C1_F3.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,1%	0,8%	0,0%	0,2%	0,2%	0,0%	0,0%	63%
2	0,3%	19,4%	0,2%	2,3%	0,9%	0,0%	0,0%	84%
3	0,0%	1,2%	0,4%	0,3%	0,5%	0,0%	0,0%	17%
4	0,1%	2,2%	0,0%	32,3%	4,2%	0,0%	0,0%	83%
5	0,1%	1,4%	0,1%	5,4%	20,9%	0,4%	0,0%	74%
6	0,0%	0,0%	0,0%	0,2%	1,3%	2,4%	0,0%	62%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	58%
Ex. Utilizador	80%	78%	52%	79%	75%	85%	92%	78%

Tabela 83: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1_F3.

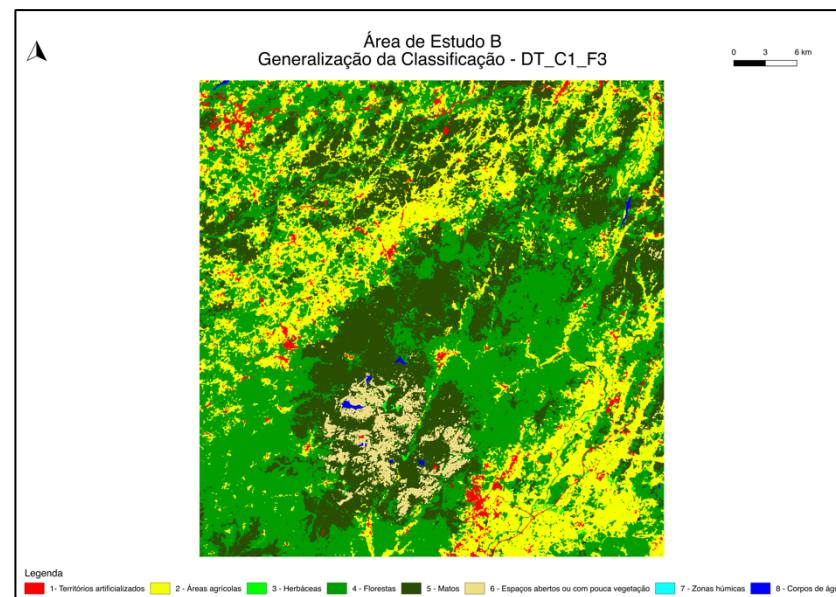


Figura 74: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C1_F3.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,0%	0,9%	0,0%	0,2%	0,1%	0,0%	0,0%	60%
2	0,2%	19,9%	0,0%	2,3%	0,6%	0,0%	0,0%	86%
3	0,0%	1,4%	0,2%	0,3%	0,5%	0,0%	0,0%	9%
4	0,1%	2,3%	0,0%	32,8%	3,5%	0,0%	0,0%	85%
5	0,1%	1,4%	0,0%	4,8%	21,6%	0,2%	0,0%	77%
6	0,0%	0,0%	0,0%	0,2%	1,3%	2,5%	0,0%	62%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	51%
Ex. Utilizador	84%	77%	80%	81%	78%	89%	94%	79%

Tabela 84: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1_F3.

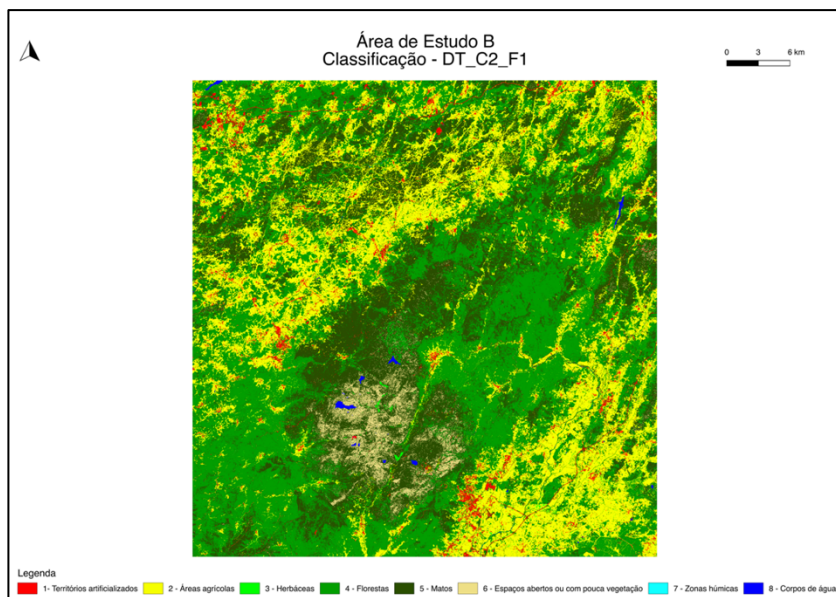


Figura 75: Área de Estudo B – MUCS obtido pela classificação com DT_C2_F1.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	1,5%	1,0%	0,0%	0,3%	0,4%	0,0%	0,0%	46%
2	0,3%	18,4%	0,2%	2,7%	1,4%	0,0%	0,0%	80%
3	0,0%	1,2%	0,2%	0,4%	0,6%	0,0%	0,0%	9%
4	0,1%	2,6%	0,0%	30,1%	5,8%	0,1%	0,0%	78%
5	0,1%	1,9%	0,1%	7,2%	18,3%	0,6%	0,0%	65%
6	0,0%	0,0%	0,0%	0,3%	2,0%	1,6%	0,0%	41%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	52%
Ex. Utilizador	72%	73%	46%	73%	64%	70%	91%	70%

Tabela 85: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C2_F1.

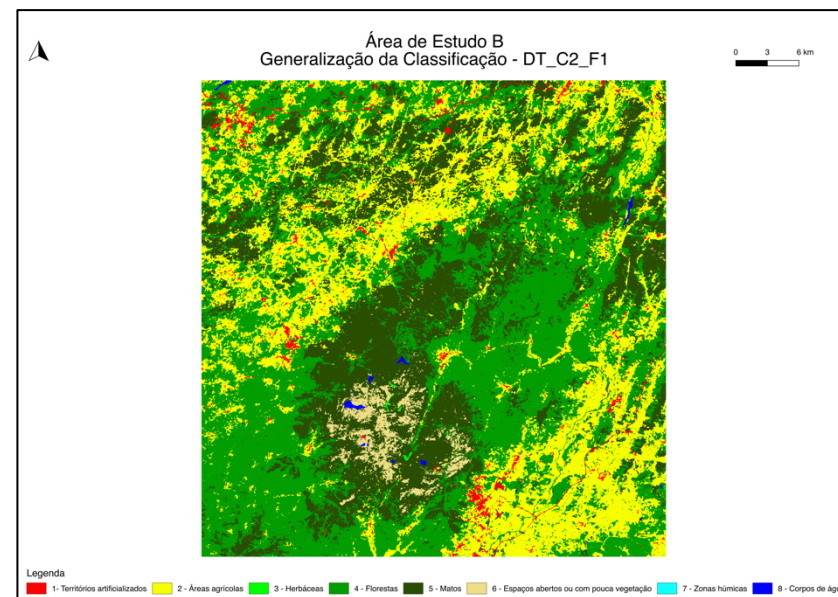


Figura 76: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C2_F1.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	1,5%	1,4%	0,0%	0,2%	0,2%	0,0%	0,0%	46%
2	0,1%	19,9%	0,0%	2,4%	0,7%	0,0%	0,0%	86%
3	0,0%	1,5%	0,1%	0,3%	0,6%	0,0%	0,0%	2%
4	0,1%	2,6%	0,0%	32,0%	4,2%	0,0%	0,0%	82%
5	0,0%	1,8%	0,0%	6,4%	19,7%	0,2%	0,0%	70%
6	0,0%	0,0%	0,0%	0,2%	2,2%	1,6%	0,0%	39%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,1%	49%
Ex. Utilizador	90%	73%	89%	77%	72%	85%	92%	75%

Tabela 86: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C2_F1.

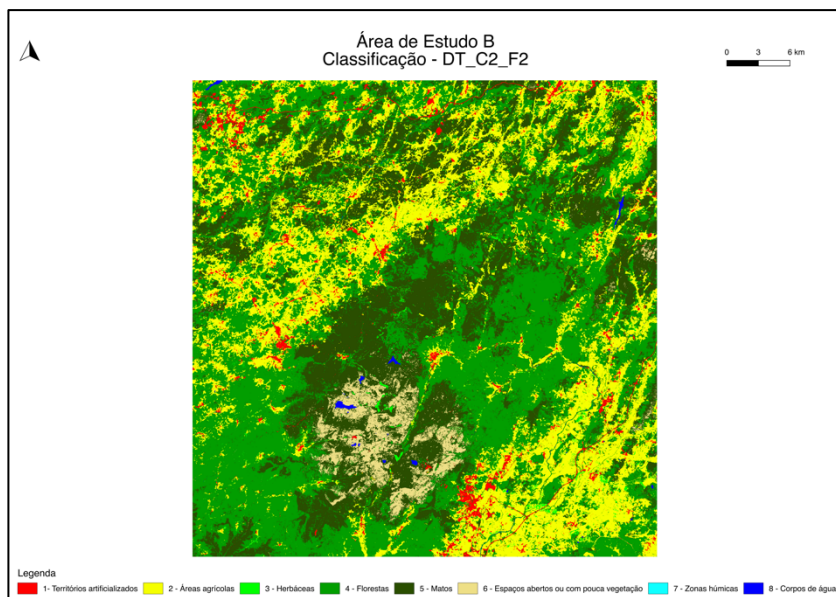


Figura 77: Área de Estudo B – MUCS obtido pela classificação com DT_C2_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,0%	0,8%	0,0%	0,2%	0,2%	0,0%	0,0%	63%
2	0,2%	19,4%	0,3%	2,2%	0,9%	0,0%	0,0%	84%
3	0,0%	1,2%	0,5%	0,3%	0,5%	0,0%	0,0%	20%
4	0,1%	2,1%	0,0%	32,6%	3,9%	0,0%	0,0%	84%
5	0,1%	1,3%	0,1%	5,0%	21,3%	0,3%	0,0%	76%
6	0,0%	0,0%	0,0%	0,2%	1,2%	2,6%	0,0%	65%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	56%
Ex. Utilizador	81%	78%	52%	80%	76%	87%	93%	79%

Tabela 87: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C2_F2.

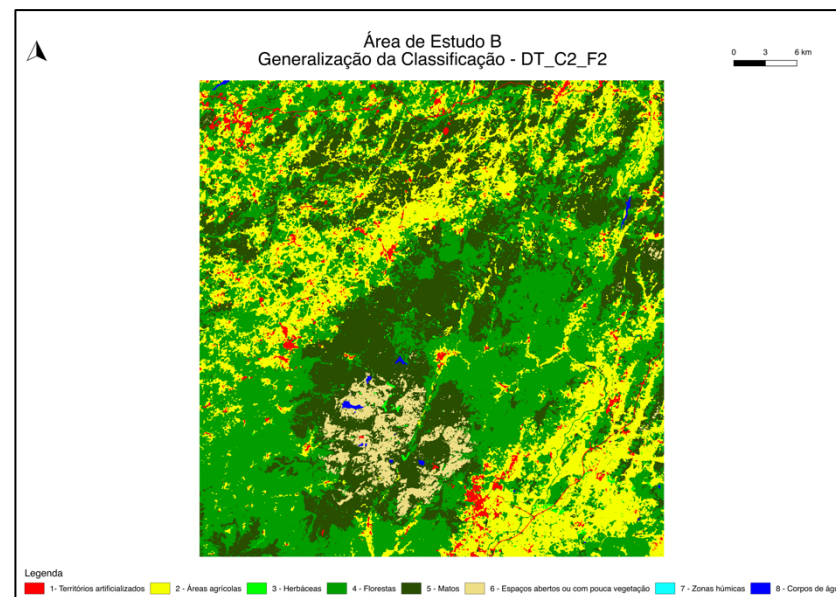


Figura 78: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C2_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,0%	0,9%	0,0%	0,2%	0,1%	0,0%	0,0%	61%
2	0,2%	19,9%	0,0%	2,2%	0,7%	0,0%	0,0%	86%
3	0,0%	1,3%	0,3%	0,3%	0,5%	0,0%	0,0%	11%
4	0,1%	2,3%	0,0%	33,2%	3,2%	0,0%	0,0%	86%
5	0,1%	1,3%	0,0%	4,4%	22,1%	0,2%	0,0%	78%
6	0,0%	0,0%	0,0%	0,2%	1,1%	2,7%	0,0%	67%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	51%
Ex. Utilizador	84%	77%	74%	82%	80%	91%	93%	80%

Tabela 88: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C2_F2.

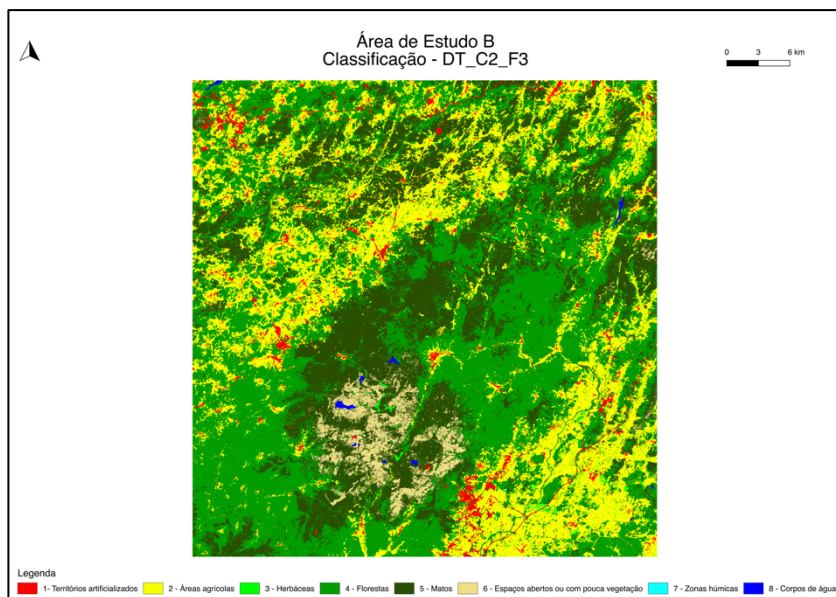


Figura 79: Área de Estudo B – MUCS obtido pela classificação com DT_C2_F3.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,0%	0,8%	0,0%	0,2%	0,2%	0,0%	0,0%	61%
2	0,3%	19,3%	0,2%	2,3%	0,9%	0,0%	0,0%	84%
3	0,0%	1,2%	0,4%	0,3%	0,5%	0,0%	0,0%	16%
4	0,1%	2,2%	0,0%	32,1%	4,3%	0,0%	0,0%	83%
5	0,1%	1,4%	0,1%	5,4%	20,8%	0,4%	0,0%	74%
6	0,0%	0,0%	0,0%	0,2%	1,3%	2,4%	0,0%	61%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	55%
Ex. Utilizador	80%	77%	53%	79%	74%	84%	92%	77%

Tabela 89: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C2_F3.

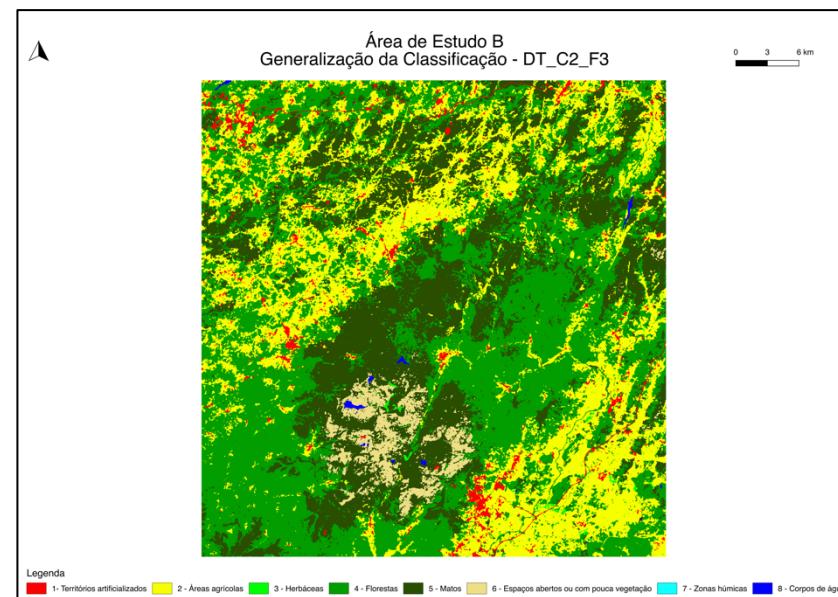


Figura 80: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C2_F3.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	1,9%	1,0%	0,0%	0,2%	0,1%	0,0%	0,0%	59%
2	0,2%	19,8%	0,0%	2,3%	0,7%	0,0%	0,0%	86%
3	0,0%	1,4%	0,2%	0,3%	0,5%	0,0%	0,0%	8%
4	0,1%	2,4%	0,0%	32,7%	3,6%	0,0%	0,0%	84%
5	0,1%	1,4%	0,0%	4,8%	21,6%	0,3%	0,0%	77%
6	0,0%	0,0%	0,0%	0,2%	1,3%	2,5%	0,0%	62%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	50%
Ex. Utilizador	84%	76%	80%	80%	77%	89%	93%	79%

Tabela 90: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C2_F3.

Anexo G – Histogramas dos Índices NDVI, NDWI e NDBI por Classe

Nas figuras que seguem neste anexo são apresentados os histogramas dos índices radiométricos NDVI, NDWI e NDBI por classe a partir das áreas de treino DT_C1.

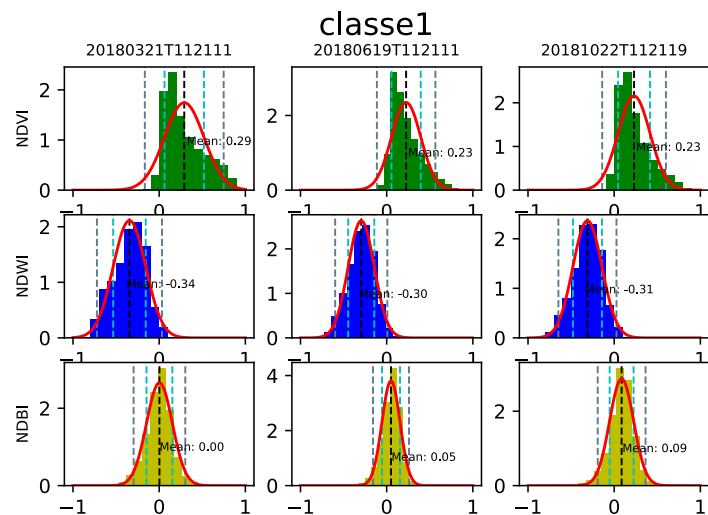


Figura 81: Área de Estudo A – Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 1.

Área de Estudo A				
Classe	Índice	Imagem	Média	Desvio Padrão
1	NDVI	20180321T112111	0,29	0,23
		20180619T112111	0,23	0,17
		20181022T112119	0,23	0,18
	NDWI	20180321T112111	-0,34	0,19
		20180619T112111	-0,30	0,15
		20181022T112119	-0,31	0,17
	NDBI	20180321T112111	0,00	0,15
		20180619T112111	0,05	0,10
		20181022T112119	0,09	0,14

Tabela 91: Área de Estudo A – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 1.

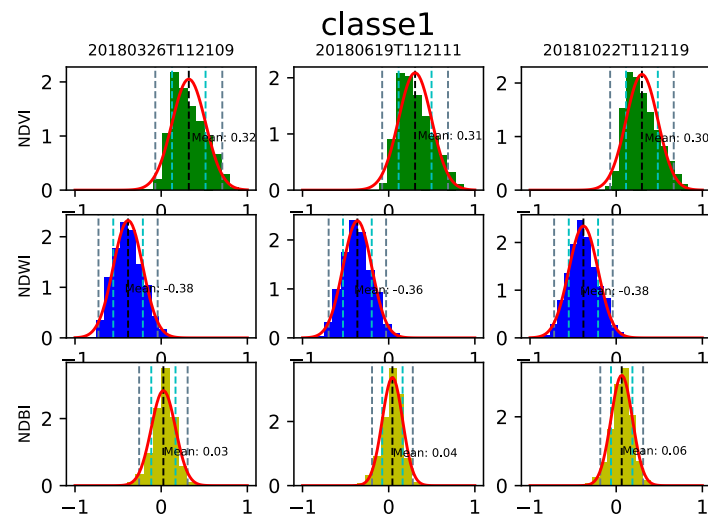


Figura 82: Área de Estudo B - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 1.

Área de Estudo B				
Classe	Índice	Imagem	Média	Desvio Padrão
1	NDVI	20180326T112109	0,32	0,19
		20180619T112111	0,31	0,19
		20181022T112119	0,30	0,18
	NDWI	20180326T112109	-0,38	0,17
		20180619T112111	-0,36	0,17
		20181022T112119	-0,38	0,17
	NDBI	20180326T112109	0,03	0,14
		20180619T112111	0,04	0,12
		20181022T112119	0,06	0,12

Tabela 92: Área de Estudo B – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 1.

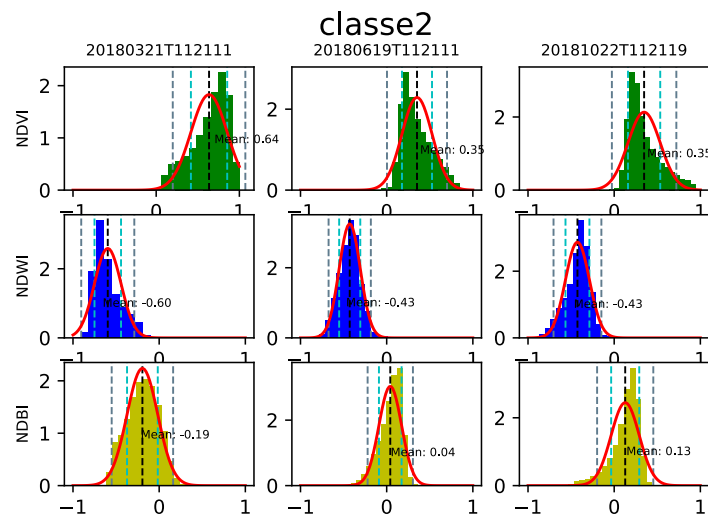


Figura 83: Área de Estudo A - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 2.

Área de Estudo A				
Classe	Índice	Imagem	Média	Desvio Padrão
2	NDVI	20180321T112111	0,64	0,22
		20180619T112111	0,35	0,17
		20181022T112119	0,35	0,19
	NDWI	20180321T112111	-0,60	0,15
		20180619T112111	-0,43	0,12
		20181022T112119	-0,43	0,14
	NDBI	20180321T112111	-0,19	0,18
		20180619T112111	0,04	0,13
		20181022T112119	0,13	0,16

Tabela 93: Área de Estudo A – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 2.

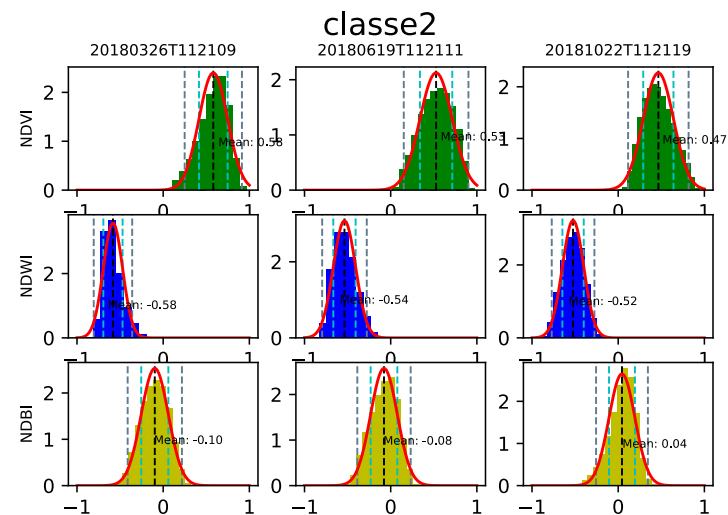


Figura 84: Área de Estudo B - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 2.

Área de Estudo B				
Classe	Índice	Imagem	Média	Desvio Padrão
2	NDVI	20180326T112109	0,58	0,17
		20180619T112111	0,53	0,19
		20181022T112119	0,47	0,18
	NDWI	20180326T112109	-0,58	0,11
		20180619T112111	-0,54	0,13
		20181022T112119	-0,52	0,12
	NDBI	20180326T112109	-0,10	0,16
		20180619T112111	-0,08	0,15
		20181022T112119	0,04	0,15

Tabela 94: Área de Estudo B – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 2.

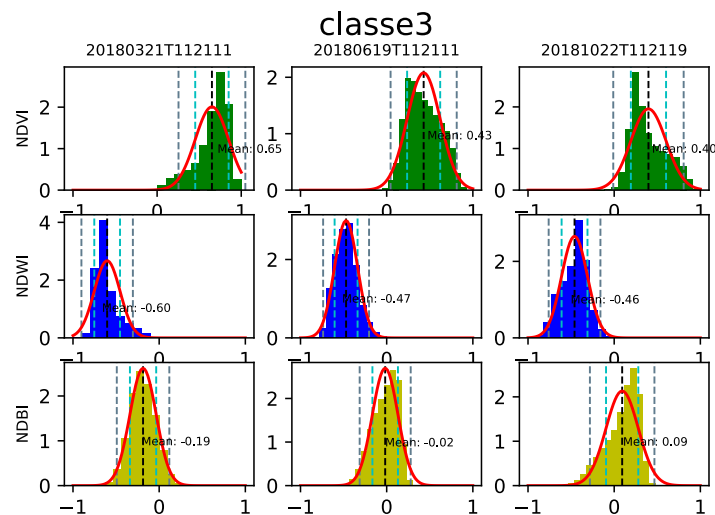


Figura 85: Área de Estudo A - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 3.

Área de Estudo A				
Classe	Índice	Imagem	Média	Desvio Padrão
3	NDVI	20180321T112111	0,65	0,20
		20180619T112111	0,43	0,19
		20181022T112119	0,40	0,20
	NDWI	20180321T112111	-0,60	0,15
		20180619T112111	-0,47	0,13
		20181022T112119	-0,46	0,15
	NDBI	20180321T112111	-0,19	0,15
		20180619T112111	-0,02	0,15
		20181022T112119	0,09	0,19

Tabela 95: Área de Estudo A – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 3.

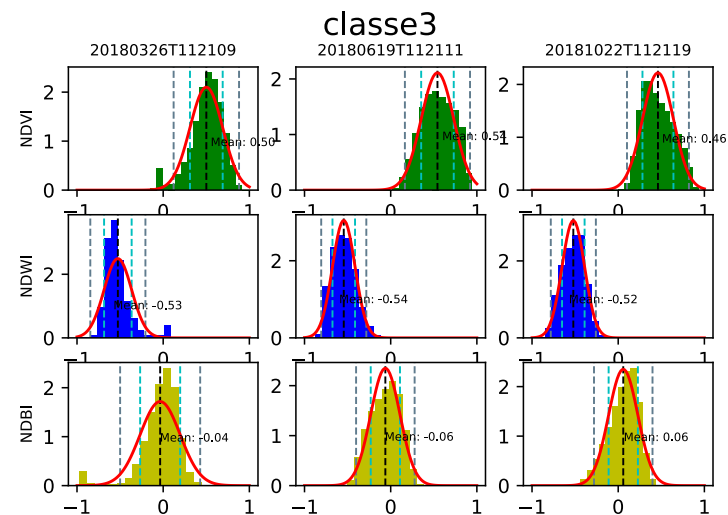


Figura 86: Área de Estudo B - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 3.

Área de Estudo B				
Classe	Índice	Imagem	Média	Desvio Padrão
3	NDVI	20180326T112109	0,50	0,19
		20180619T112111	0,54	0,19
		20181022T112119	0,46	0,18
	NDWI	20180326T112109	-0,53	0,16
		20180619T112111	-0,54	0,13
		20181022T112119	-0,52	0,13
	NDBI	20180326T112109	-0,04	0,23
		20180619T112111	-0,06	0,17
		20181022T112119	0,06	0,17

Tabela 96: Área de Estudo B – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 3.

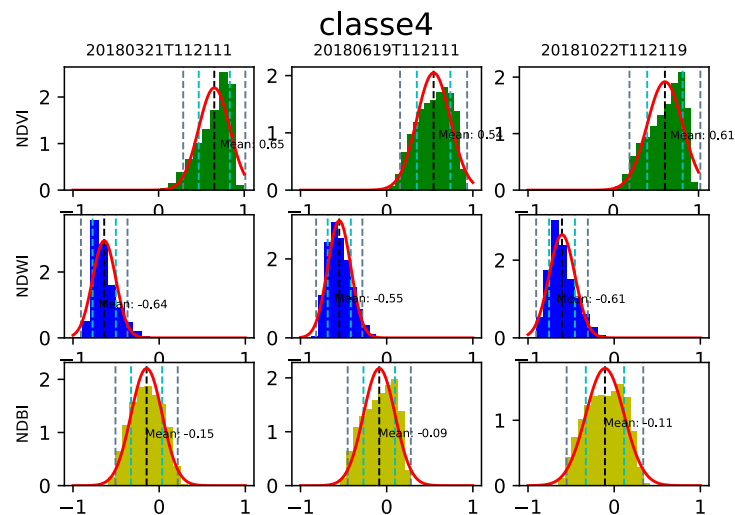


Figura 87: Área de Estudo A - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 4.

Área de Estudo A				
Classe	Índice	Imagem	Média	Desvio Padrão
4	NDVI	20180321T112111	0,65	0,18
		20180619T112111	0,54	0,19
		20181022T112119	0,61	0,21
	NDWI	20180321T112111	-0,64	0,14
		20180619T112111	-0,55	0,13
		20181022T112119	-0,61	0,15
	NDBI	20180321T112111	-0,15	0,18
		20180619T112111	-0,09	0,18
		20181022T112119	-0,11	0,22

Tabela 97: Área de Estudo A – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 4.

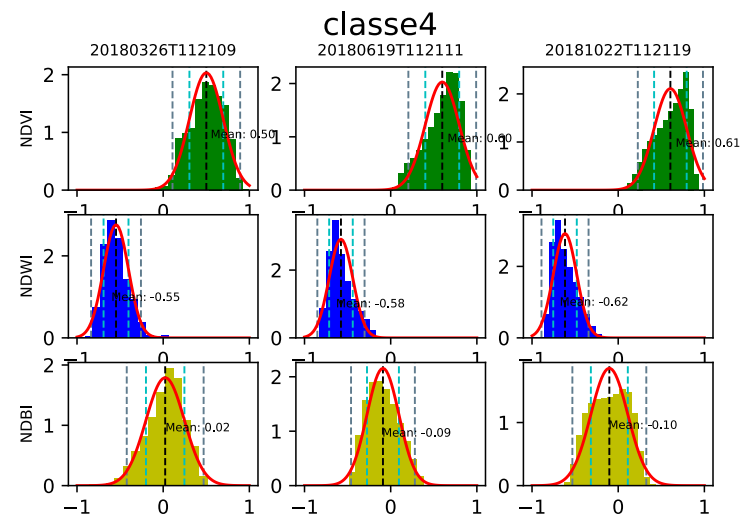


Figura 88: Área de Estudo B - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 4.

Área de Estudo B				
Classe	Índice	Imagem	Média	Desvio Padrão
4	NDVI	20180326T112109	0,50	0,20
		20180619T112111	0,60	0,20
		20181022T112119	0,61	0,19
	NDWI	20180326T112109	-0,55	0,14
		20180619T112111	-0,58	0,14
		20181022T112119	-0,62	0,14
	NDBI	20180326T112109	0,02	0,22
		20180619T112111	-0,09	0,19
		20181022T112119	-0,10	0,21

Tabela 98: Área de Estudo B – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 4.

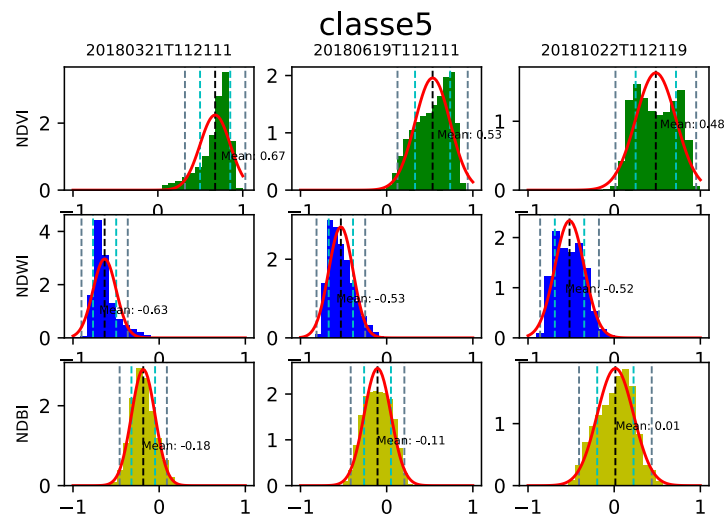


Figura 89: Área de Estudo A - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 5.

Área de Estudo A				
Classe	Índice	Imagem	Média	Desvio Padrão
5	NDVI	20180321T112111	0,67	0,18
		20180619T112111	0,53	0,20
		20181022T112119	0,48	0,23
	NDWI	20180321T112111	-0,63	0,13
		20180619T112111	-0,53	0,14
		20181022T112119	-0,52	0,17
	NDBI	20180321T112111	-0,18	0,14
		20180619T112111	-0,11	0,16
		20181022T112119	0,01	0,21

Tabela 99: Área de Estudo A – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 5.

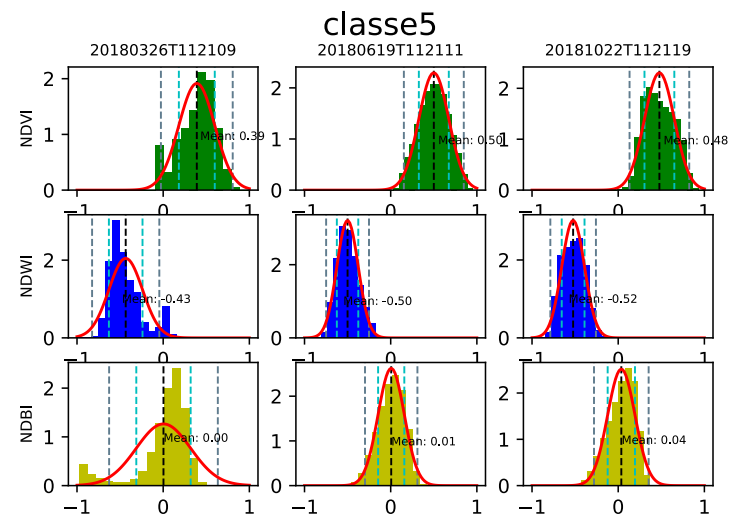


Figura 90: Área de Estudo B - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 5.

Área de Estudo B				
Classe	Índice	Imagem	Média	Desvio Padrão
5	NDVI	20180326T112109	0,39	0,21
		20180619T112111	0,50	0,17
		20181022T112119	0,48	0,17
	NDWI	20180326T112109	-0,43	0,20
		20180619T112111	-0,50	0,12
		20181022T112119	-0,52	0,13
	NDBI	20180326T112109	0,00	0,32
		20180619T112111	0,01	0,15
		20181022T112119	0,04	0,16

Tabela 100: Área de Estudo B – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 5.

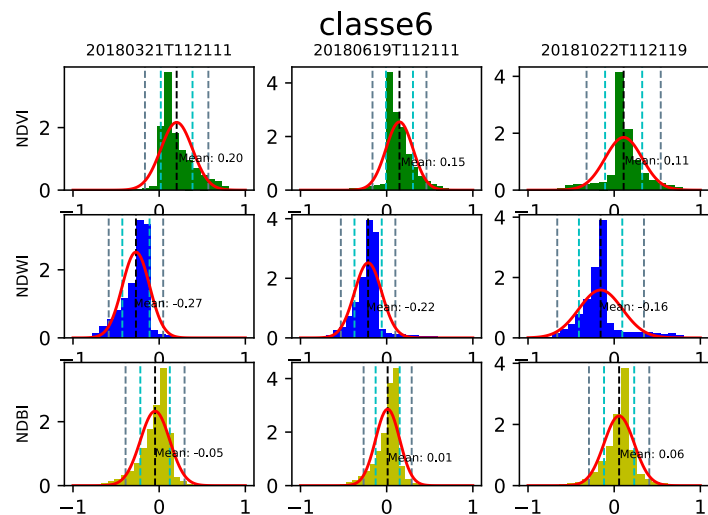


Figura 91: Área de Estudo A - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 6.

Área de Estudo A				
Classe	Índice	Imagem	Média	Desvio Padrão
6	NDVI	20180321T112111	0,20	0,18
		20180619T112111	0,15	0,16
		20181022T112119	0,11	0,22
	NDWI	20180321T112111	-0,27	0,16
		20180619T112111	-0,22	0,16
		20181022T112119	-0,16	0,25
	NDBI	20180321T112111	-0,05	0,17
		20180619T112111	0,01	0,14
		20181022T112119	0,06	0,17

Tabela 101: Área de Estudo A – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 6.

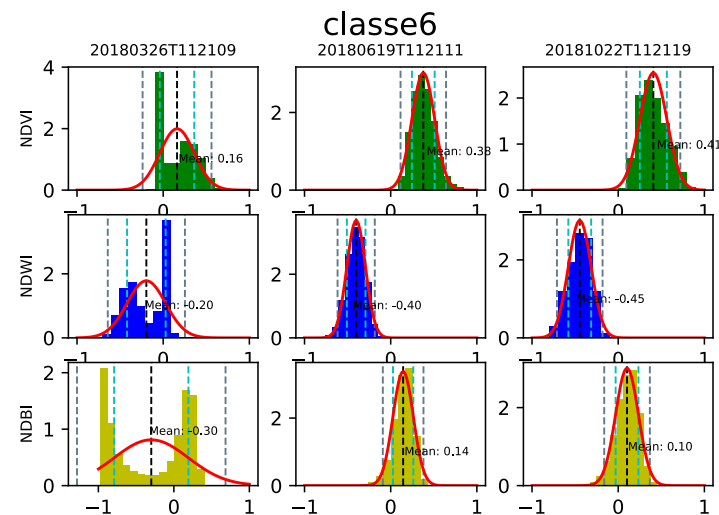


Figura 92: Área de Estudo B - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 6.

Área de Estudo B				
Classe	Índice	Imagem	Média	Desvio Padrão
6	NDVI	20180326T112109	0,39	0,21
		20180619T112111	0,50	0,17
		20181022T112119	0,48	0,17
	NDWI	20180326T112109	-0,43	0,20
		20180619T112111	-0,50	0,12
		20181022T112119	-0,52	0,13
	NDBI	20180326T112109	0,00	0,43
		20180619T112111	0,01	0,15
		20181022T112119	0,04	0,16

Tabela 102: Área de Estudo B – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 6.

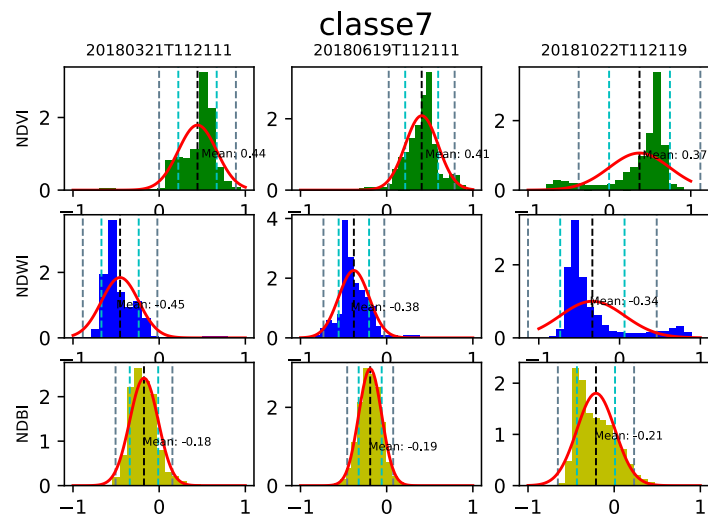


Figura 93: Área de Estudo A - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 7.

Área de Estudo A				
Classe	Índice	Imagem	Média	Desvio Padrão
7	NDVI	20180321T112111	0,44	0,22
		20180619T112111	0,41	0,19
		20181022T112119	0,37	0,37
	NDWI	20180321T112111	-0,45	0,22
		20180619T112111	-0,38	0,18
		20181022T112119	-0,34	0,40
	NDBI	20180321T112111	-0,18	0,16
		20180619T112111	-0,19	0,13
		20181022T112119	-0,21	0,22

Tabela 103: Área de Estudo A – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 7.

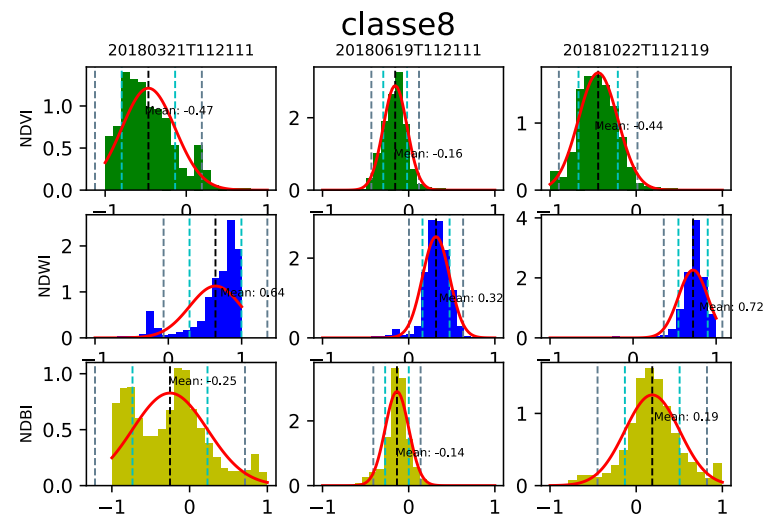


Figura 94: Área de Estudo A - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 8.

Área de Estudo A				
Classe	Índice	Imagem	Média	Desvio Padrão
8	NDVI	20180321T112111	-0,47	0,33
		20180619T112111	-0,16	0,14
		20181022T112119	-0,44	0,23
	NDWI	20180321T112111	0,64	0,35
		20180619T112111	0,32	0,16
		20181022T112119	0,72	0,18
	NDBI	20180321T112111	-0,25	0,48
		20180619T112111	-0,14	0,14
		20181022T112119	0,19	0,32

Tabela 104: Área de Estudo A – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 8.

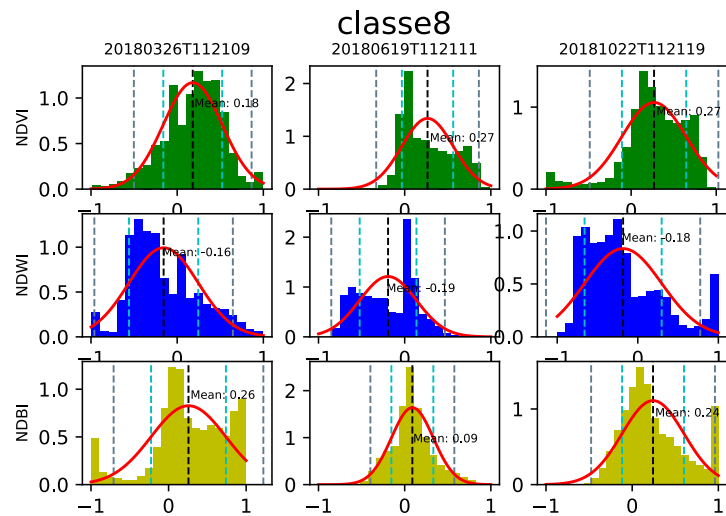


Figura 95: Área de Estudo B - Histogramas dos índices radiométricos NDVI, NDWI e NDBI para a classe 8.

Área de Estudo B				
Classe	Índice	Imagem	Média	Desvio Padrão
8	NDVI	20180326T112109	0,18	0,34
		20180619T112111	0,27	0,30
		20181022T112119	0,27	0,38
	NDWI	20180326T112109	-0,16	0,40
		20180619T112111	-0,19	0,33
		20181022T112119	-0,18	0,48
	NDBI	20180326T112109	0,26	0,48
		20180619T112111	0,09	0,24
		20181022T112119	0,24	0,36

Tabela 105: Área de Estudo B – Médias e desvios padrões para os índices radiométricos NDVI, NDWI e NDBI para a classe 8.

Anexo H – Classificações e Respetivas Generalizações a Partir dos Dados de Treino Obtidos pela Filtragem dos Dados de Treino Extraídos da COS 2015

De seguida, apresentam-se os MUCS obtidos com os dados de treino obtidos pela filtragem dos dados de treino extraídos da COS 2015 e as respetivas matrizes de confusão.

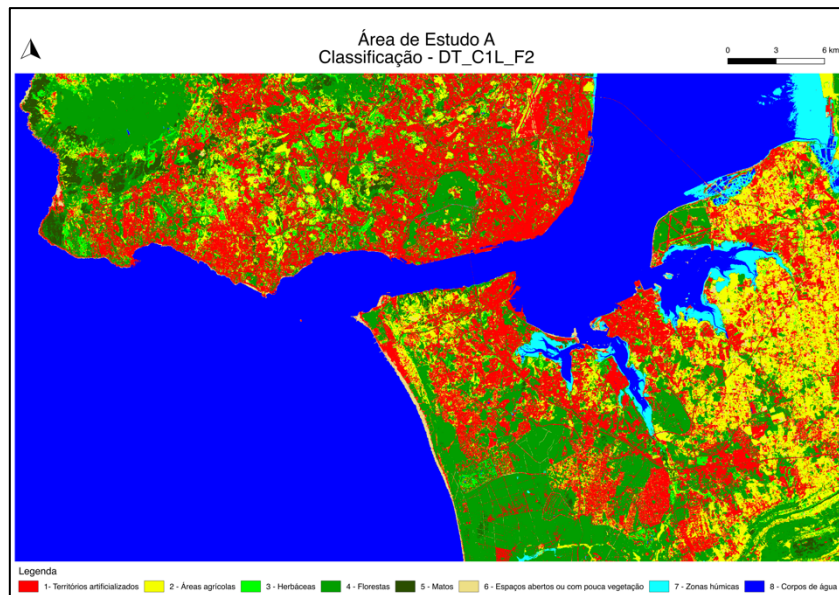


Figura 96: Área de Estudo A – MUCS obtido pela classificação com DT_C1L_F2.

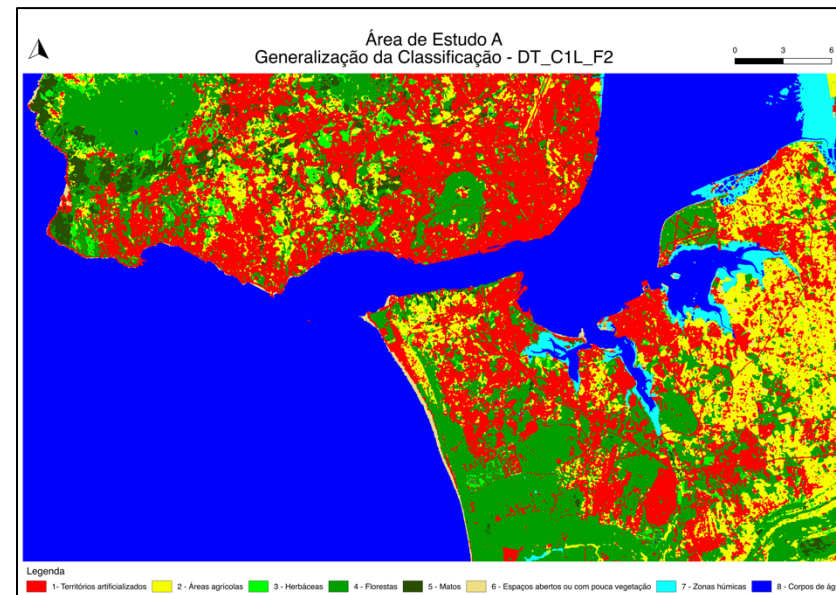


Figura 97: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C1L_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	17,3%	2,3%	1,0%	2,5%	0,6%	0,5%	0,0%	0,1%	71%
2	1,3%	5,9%	0,6%	0,8%	0,2%	0,4%	0,0%	0,0%	64%
3	0,4%	1,0%	1,1%	0,5%	0,2%	0,1%	0,0%	0,0%	34%
4	0,9%	0,6%	0,2%	11,7%	0,3%	0,1%	0,0%	0,0%	85%
5	0,3%	0,6%	0,4%	1,0%	1,8%	0,1%	0,0%	0,0%	43%
6	0,1%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,0%	63%
7	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,7%	0,0%	88%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,1%	1,1%	42,8%	97%
Ex. Utilizador	85%	57%	32%	71%	60%	16%	37%	100%	82%

Tabela 106: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1L_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	19,6%	1,8%	0,3%	2,1%	0,2%	0,1%	0,0%	0,1%	81%
2	1,3%	6,6%	0,3%	0,7%	0,1%	0,2%	0,0%	0,0%	72%
3	0,5%	1,0%	1,0%	0,5%	0,1%	0,0%	0,0%	0,0%	32%
4	0,9%	0,5%	0,1%	12,2%	0,1%	0,0%	0,0%	0,0%	89%
5	0,5%	0,6%	0,3%	1,0%	1,8%	0,0%	0,0%	0,0%	42%
6	0,1%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,0%	60%
7	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,7%	0,0%	91%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,1%	1,2%	42,8%	97%
Ex. Utilizador	85%	62%	50%	73%	78%	36%	37%	100%	85%

Tabela 107: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1L_F2.

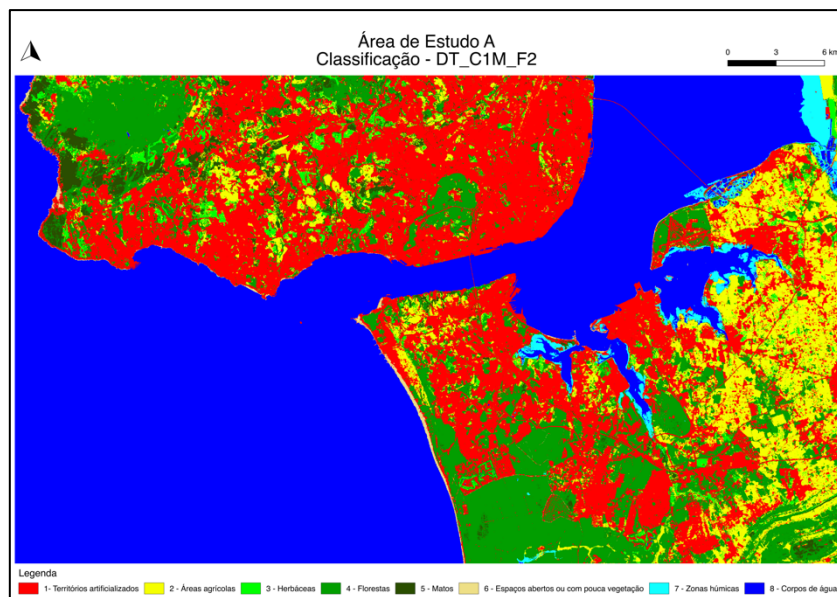


Figura 98: Área de Estudo A – MUCS obtido pela classificação com DT_C1M_F2.

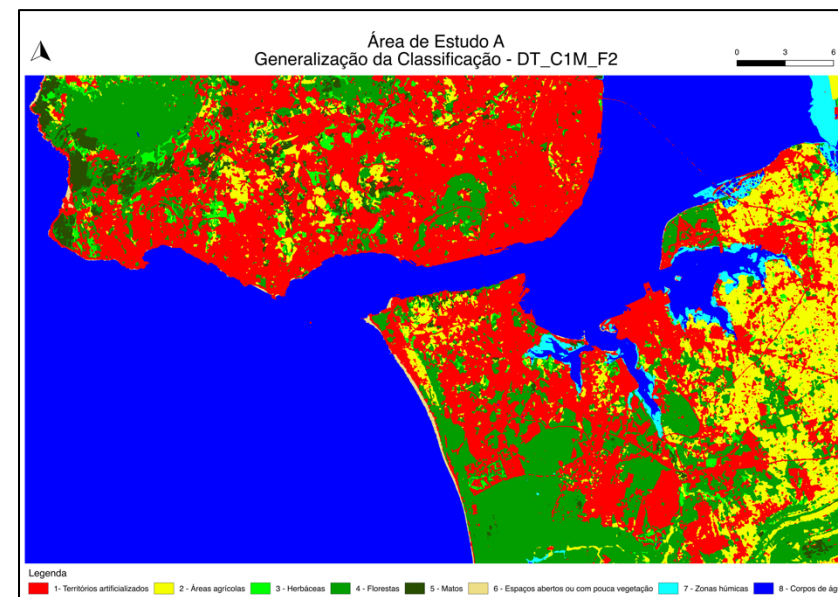


Figura 99: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C1M_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	22,4%	0,7%	0,2%	0,9%	0,1%	0,0%	0,0%	0,0%	92%
2	1,2%	6,8%	0,5%	0,6%	0,1%	0,0%	0,0%	0,0%	74%
3	0,7%	0,7%	1,3%	0,3%	0,1%	0,0%	0,0%	0,0%	41%
4	1,0%	0,4%	0,1%	11,9%	0,2%	0,0%	0,0%	0,0%	87%
5	0,8%	0,4%	0,3%	0,8%	1,9%	0,0%	0,0%	0,0%	45%
6	0,1%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,0%	57%
7	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,7%	0,0%	85%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	43,3%	98%
Ex. Utilizador	85%	75%	54%	82%	77%	74%	52%	100%	89%

Tabela 108: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1M_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	22,8%	0,6%	0,1%	0,8%	0,0%	0,0%	0,0%	0,0%	94%
2	1,2%	7,1%	0,3%	0,5%	0,1%	0,0%	0,0%	0,0%	77%
3	0,9%	0,7%	1,2%	0,3%	0,0%	0,0%	0,0%	0,0%	37%
4	1,1%	0,4%	0,1%	12,1%	0,1%	0,0%	0,0%	0,0%	88%
5	1,1%	0,4%	0,2%	0,8%	1,8%	0,0%	0,0%	0,0%	42%
6	0,1%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,0%	51%
7	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,7%	0,0%	87%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	43,3%	98%
Ex. Utilizador	83%	78%	65%	83%	87%	82%	52%	100%	89%

Tabela 109: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1M_F2.

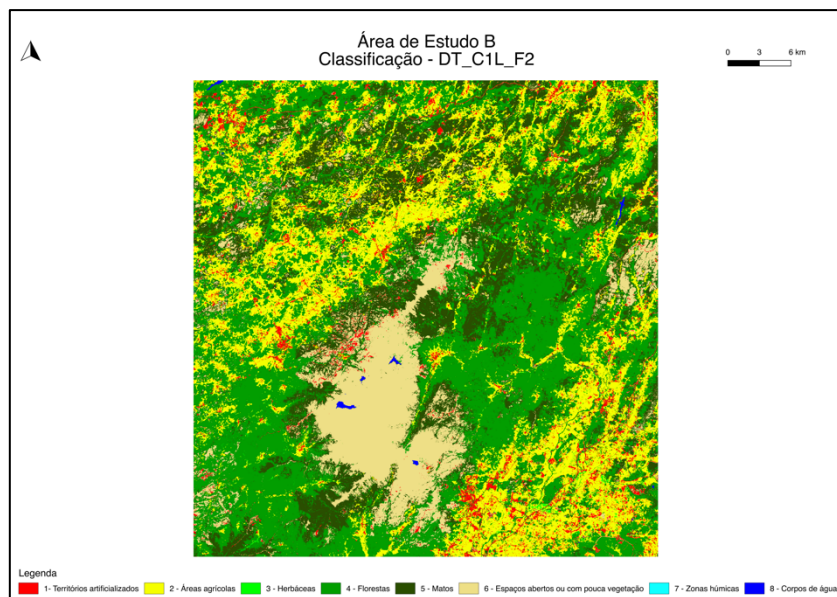


Figura 100: Área de Estudo B – MUCS obtido pela classificação com DT_C1L_F2.

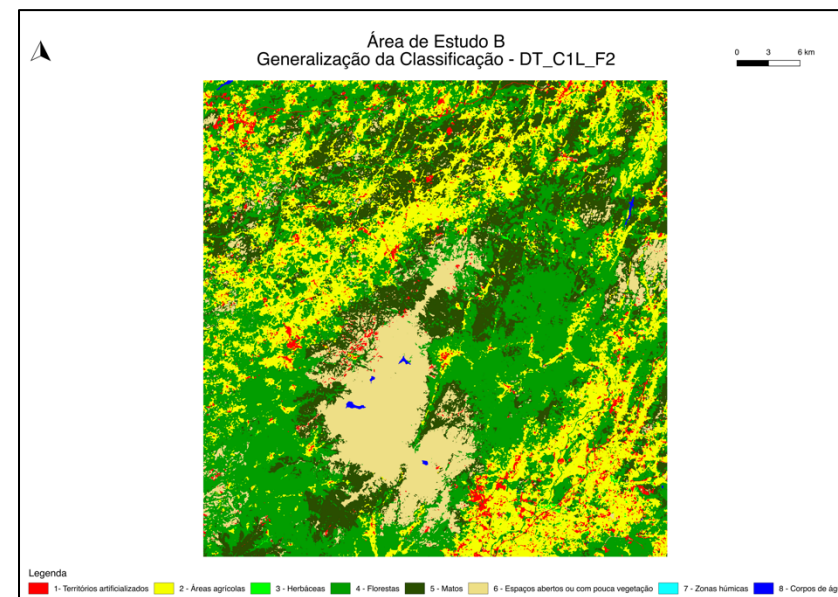


Figura 101: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C1L_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	1,7%	1,0%	0,0%	0,2%	0,2%	0,1%	0,0%	53%
2	1,2%	18,2%	0,3%	2,2%	1,0%	0,2%	0,0%	79%
3	0,1%	1,1%	0,3%	0,3%	0,4%	0,2%	0,0%	11%
4	0,4%	2,1%	0,0%	30,4%	3,8%	2,0%	0,0%	78%
5	0,5%	1,3%	0,1%	4,9%	15,2%	6,2%	0,0%	54%
6	0,0%	0,0%	0,0%	0,1%	0,4%	3,4%	0,0%	85%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,1%	37%
Ex. Utilizador	43%	77%	41%	80%	72%	28%	95%	69%

Tabela 110: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1L_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	1,7%	1,2%	0,0%	0,2%	0,1%	0,1%	0,0%	52%
2	0,8%	19,1%	0,0%	2,2%	0,8%	0,1%	0,0%	83%
3	0,1%	1,3%	0,1%	0,3%	0,4%	0,2%	0,0%	4%
4	0,3%	2,2%	0,0%	31,3%	3,0%	1,9%	0,0%	81%
5	0,4%	1,4%	0,0%	4,5%	15,9%	6,1%	0,0%	56%
6	0,0%	0,0%	0,0%	0,1%	0,4%	3,4%	0,0%	87%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,1%	38%
Ex. Utilizador	51%	76%	68%	81%	77%	29%	96%	72%

Tabela 111: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1L_F2.

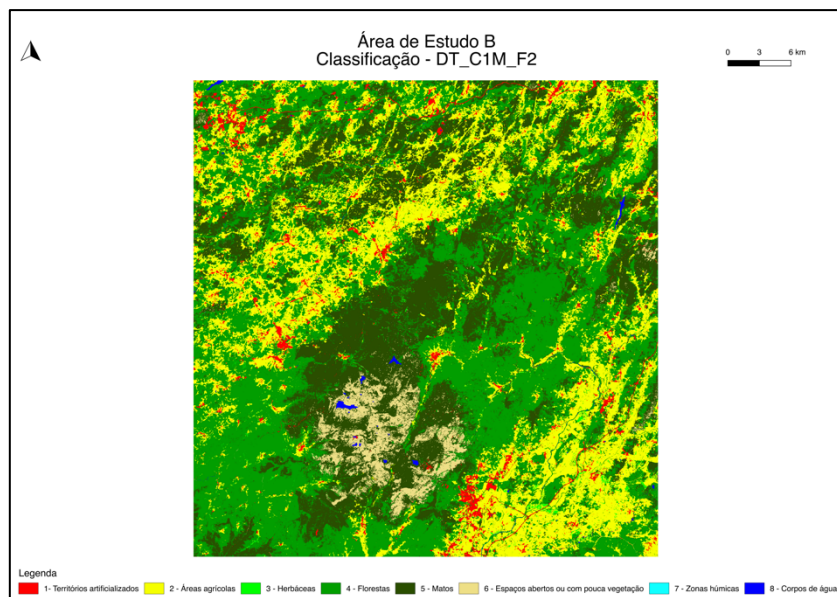


Figura 102: Área de Estudo B – MUCS obtido pela classificação com DT_C1M_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,1%	0,7%	0,0%	0,2%	0,2%	0,0%	0,0%	66%
2	0,3%	19,4%	0,3%	2,2%	0,8%	0,0%	0,0%	84%
3	0,0%	1,1%	0,5%	0,3%	0,5%	0,0%	0,0%	18%
4	0,1%	2,1%	0,1%	32,4%	4,1%	0,0%	0,0%	84%
5	0,1%	1,3%	0,1%	4,9%	21,5%	0,3%	0,0%	76%
6	0,0%	0,0%	0,0%	0,2%	1,2%	2,6%	0,0%	65%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	61%
Ex. Utilizador	79%	79%	48%	81%	76%	86%	86%	79%

Tabela 112: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1M_F2.

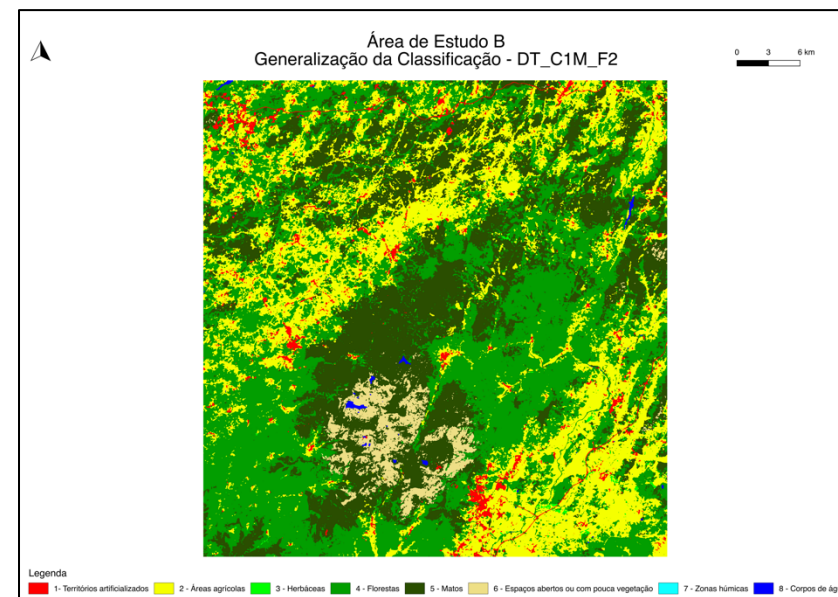


Figura 103: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C1M_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,1%	0,8%	0,0%	0,2%	0,1%	0,0%	0,0%	64%
2	0,2%	20,0%	0,1%	2,2%	0,6%	0,0%	0,0%	87%
3	0,0%	1,3%	0,2%	0,3%	0,6%	0,0%	0,0%	9%
4	0,1%	2,2%	0,0%	33,0%	3,4%	0,0%	0,0%	85%
5	0,1%	1,3%	0,0%	4,4%	22,2%	0,2%	0,0%	79%
6	0,0%	0,0%	0,0%	0,1%	1,1%	2,7%	0,0%	67%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	53%
Ex. Utilizador	82%	78%	70%	82%	79%	91%	88%	80%

Tabela 113: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1M_F2.

Anexo I – Classificações e Respetivas Generalizações a Partir dos Dados de Treino Extraídos do OSM

Neste anexo apresentam-se os MUCS obtidos com os dados de treino extraídos do OSM e as respetivas matrizes de confusão.

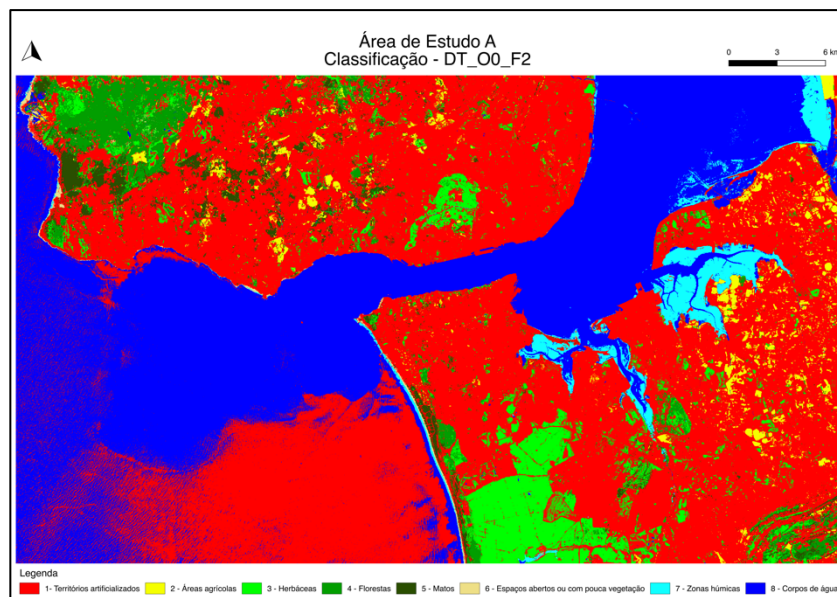


Figura 104: Área de Estudo A – MUCS obtido pela classificação com DT_O0_F2.

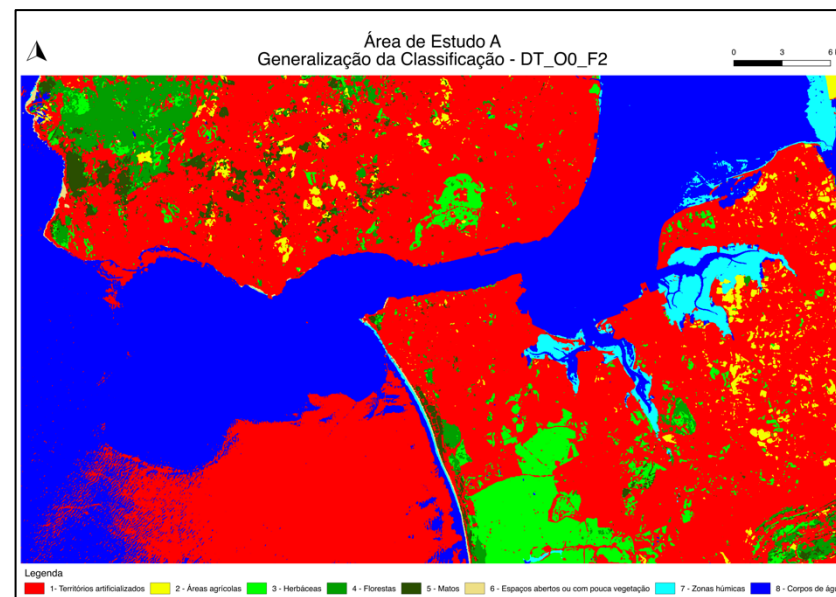


Figura 105: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_O0_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	23,8%	0,0%	0,2%	0,0%	0,1%	0,0%	0,0%	0,1%	98%
2	6,9%	1,7%	0,2%	0,1%	0,3%	0,0%	0,0%	0,0%	18%
3	2,3%	0,2%	0,4%	0,1%	0,2%	0,0%	0,0%	0,0%	12%
4	5,6%	0,1%	4,7%	2,8%	0,6%	0,0%	0,0%	0,0%	20%
5	2,4%	0,1%	0,3%	0,5%	0,9%	0,0%	0,0%	0,1%	21%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	33%
7	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	79%
8	11,8%	0,0%	0,0%	0,0%	0,0%	0,1%	1,6%	30,7%	69%
Ex. Utilizador	45%	79%	6%	81%	42%	59%	27%	99%	61%

Tabela 114: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_O0_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	24,0%	0,0%	0,1%	0,0%	0,0%	0,0%	0,0%	0,1%	99%
2	7,3%	1,5%	0,1%	0,1%	0,2%	0,0%	0,0%	0,0%	16%
3	2,4%	0,1%	0,3%	0,0%	0,2%	0,0%	0,0%	0,0%	10%
4	6,0%	0,1%	4,5%	2,8%	0,4%	0,0%	0,0%	0,0%	20%
5	2,6%	0,1%	0,3%	0,5%	0,8%	0,0%	0,0%	0,0%	19%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	32%
7	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,7%	0,1%	82%
8	11,3%	0,0%	0,0%	0,0%	0,0%	0,0%	1,6%	31,3%	71%
Ex. Utilizador	45%	83%	6%	82%	47%	68%	28%	99%	61%

Tabela 115: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_O0_F2.

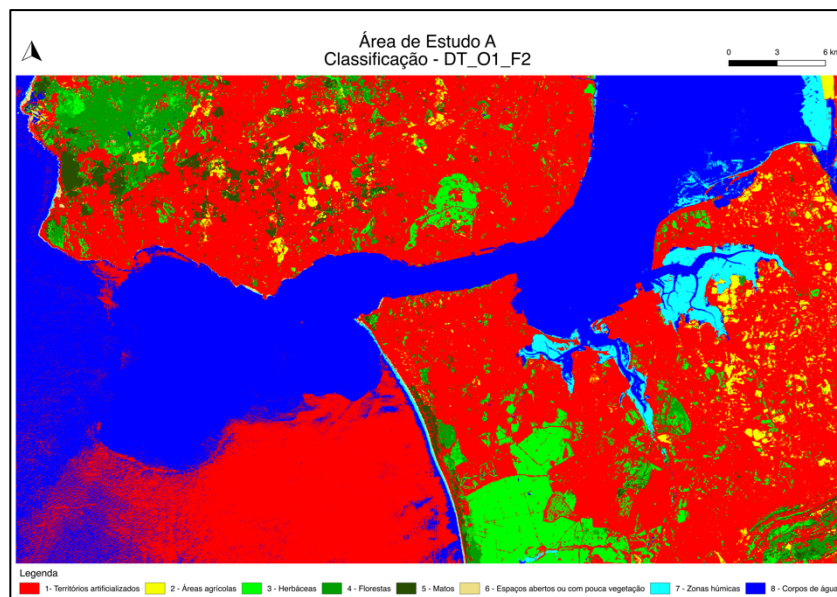


Figura 106: Área de Estudo A – MUCS obtido pela classificação com DT_O1_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	23,8%	0,0%	0,2%	0,0%	0,1%	0,0%	0,0%	0,1%	98%
2	6,9%	1,7%	0,2%	0,1%	0,3%	0,0%	0,0%	0,0%	18%
3	2,3%	0,2%	0,4%	0,1%	0,2%	0,0%	0,0%	0,0%	12%
4	5,6%	0,1%	4,7%	2,8%	0,6%	0,0%	0,0%	0,0%	20%
5	2,4%	0,1%	0,3%	0,5%	0,9%	0,0%	0,0%	0,1%	21%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	32%
7	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	79%
8	12,1%	0,0%	0,0%	0,0%	0,0%	0,1%	1,6%	30,4%	69%
Ex. Utilizador	45%	79%	6%	81%	42%	58%	27%	99%	61%

Tabela 116: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_O1_F2.

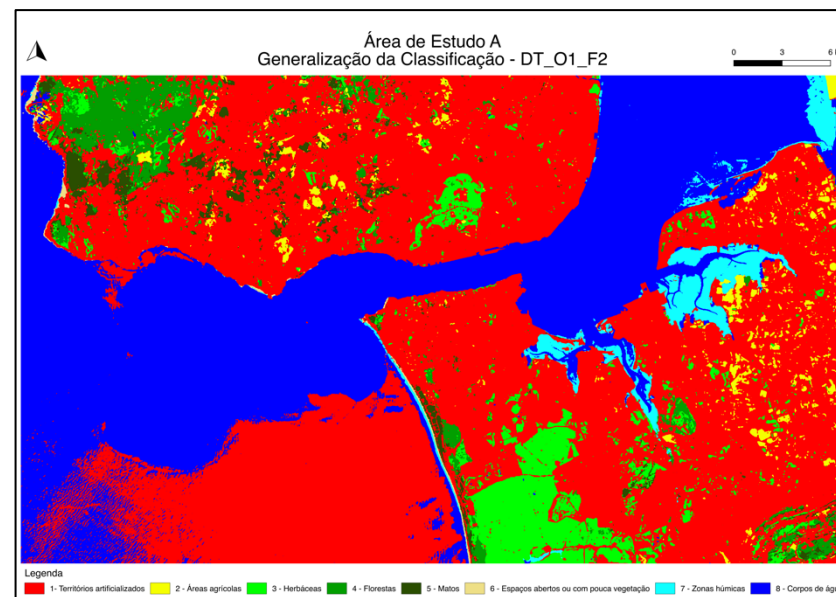


Figura 107: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_O1_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	24,0%	0,0%	0,1%	0,0%	0,0%	0,0%	0,0%	0,1%	99%
2	7,3%	1,5%	0,1%	0,1%	0,2%	0,0%	0,0%	0,0%	16%
3	2,4%	0,1%	0,3%	0,0%	0,2%	0,0%	0,0%	0,0%	9%
4	6,0%	0,1%	4,5%	2,7%	0,4%	0,0%	0,0%	0,0%	20%
5	2,6%	0,1%	0,3%	0,5%	0,8%	0,0%	0,0%	0,0%	19%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	31%
7	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	82%
8	12,0%	0,0%	0,0%	0,0%	0,0%	0,0%	1,6%	30,5%	69%
Ex. Utilizador	44%	83%	6%	82%	47%	67%	28%	99%	61%

Tabela 117: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_O1_F2.

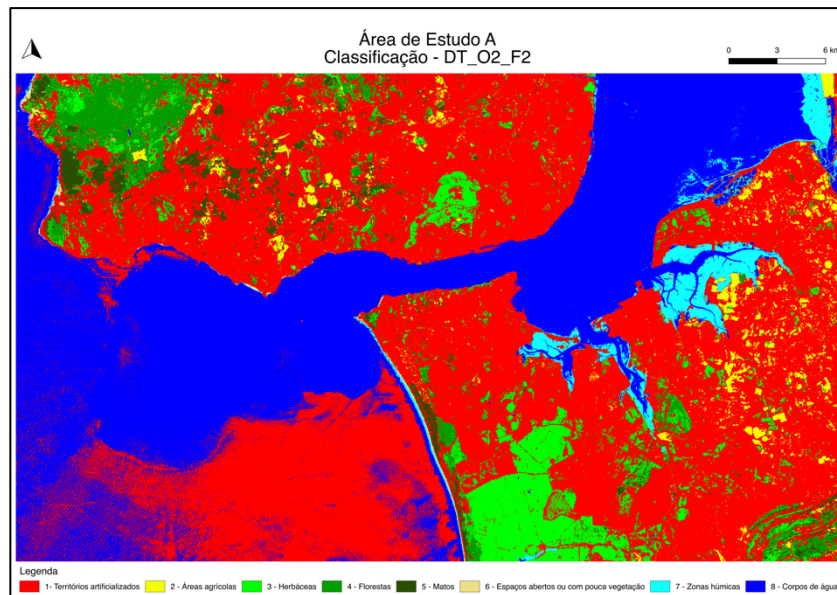


Figura 108: Área de Estudo A – MUCS obtido pela classificação com DT_O2_F2.

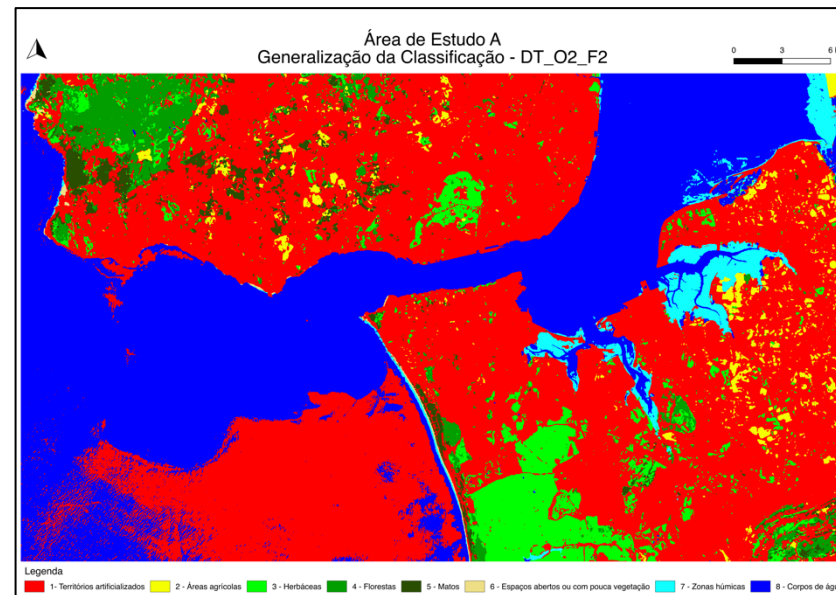


Figura 109: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_O2_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	23,8%	0,0%	0,3%	0,0%	0,1%	0,0%	0,0%	0,1%	98%
2	6,7%	1,7%	0,3%	0,1%	0,4%	0,0%	0,0%	0,0%	19%
3	2,2%	0,2%	0,4%	0,1%	0,3%	0,0%	0,0%	0,0%	13%
4	5,3%	0,1%	4,8%	2,9%	0,6%	0,0%	0,0%	0,0%	21%
5	2,3%	0,1%	0,4%	0,5%	1,0%	0,0%	0,0%	0,0%	23%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	29%
7	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	80%
8	11,3%	0,0%	0,0%	0,0%	0,0%	0,0%	1,6%	31,2%	71%
Ex. Utilizador	46%	81%	6%	81%	42%	68%	27%	99%	62%

Tabela 118: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_O2_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	24,0%	0,0%	0,2%	0,0%	0,0%	0,0%	0,0%	0,1%	99%
2	7,1%	1,5%	0,1%	0,1%	0,3%	0,0%	0,0%	0,0%	17%
3	2,4%	0,1%	0,3%	0,1%	0,2%	0,0%	0,0%	0,0%	10%
4	5,7%	0,1%	4,7%	2,8%	0,5%	0,0%	0,0%	0,0%	20%
5	2,5%	0,1%	0,3%	0,5%	0,9%	0,0%	0,0%	0,0%	21%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	28%
7	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,7%	0,0%	82%
8	11,1%	0,0%	0,0%	0,0%	0,0%	0,0%	1,6%	31,5%	71%
Ex. Utilizador	45%	84%	6%	81%	46%	76%	28%	99%	62%

Tabela 119: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_O2_F2.

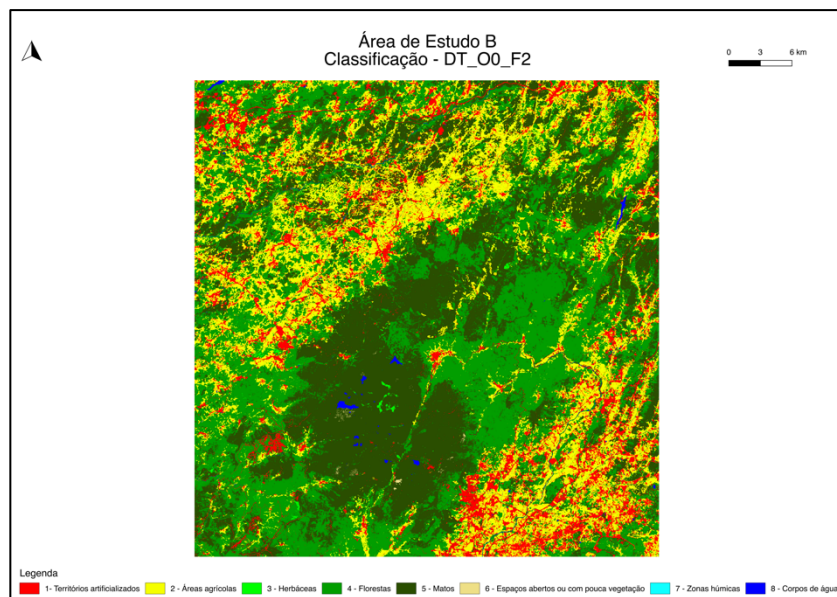


Figura 110: Área de Estudo B – MUCS obtido pela classificação com DT_O0_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,8%	0,2%	0,0%	0,1%	0,2%	0,0%	0,0%	86%
2	4,2%	16,0%	0,0%	1,9%	0,8%	0,0%	0,0%	70%
3	0,3%	1,2%	0,2%	0,2%	0,5%	0,0%	0,0%	8%
4	1,4%	2,4%	0,1%	28,0%	6,8%	0,0%	0,1%	72%
5	1,0%	2,0%	0,0%	4,5%	20,6%	0,0%	0,0%	73%
6	0,0%	0,0%	0,0%	0,1%	3,8%	0,0%	0,0%	1%
8	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,2%	64%
Ex. Utilizador	29%	73%	61%	80%	63%	95%	55%	68%

Tabela 120: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_O0_F2.

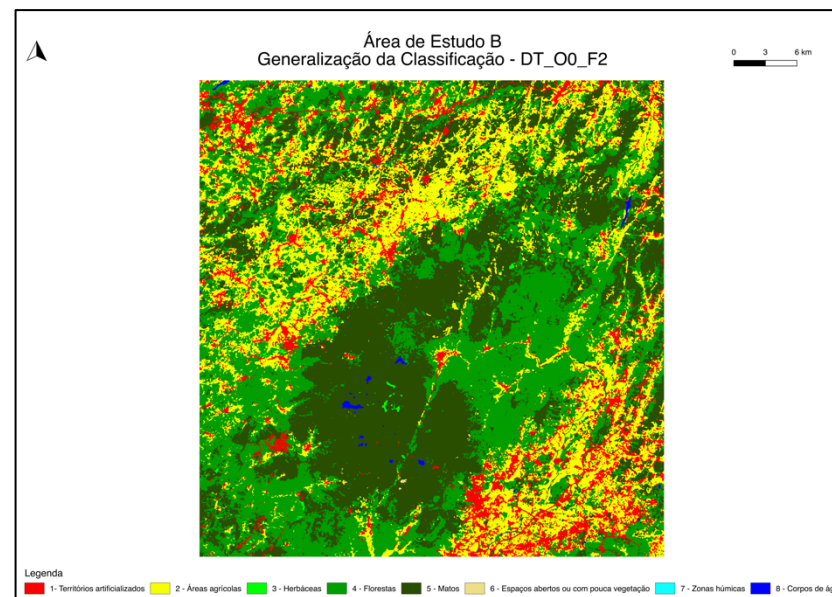


Figura 111: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_O0_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,6%	0,4%	0,0%	0,1%	0,1%	0,0%	0,0%	79%
2	1,5%	18,4%	0,0%	2,4%	0,8%	0,0%	0,0%	80%
3	0,1%	1,4%	0,1%	0,3%	0,6%	0,0%	0,0%	4%
4	0,4%	2,5%	0,0%	30,6%	5,2%	0,0%	0,0%	79%
5	0,3%	2,2%	0,0%	5,1%	20,6%	0,0%	0,0%	73%
6	0,0%	0,0%	0,0%	0,3%	3,6%	0,1%	0,0%	2%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	55%
Ex. Utilizador	52%	74%	94%	79%	67%	99%	83%	73%

Tabela 121: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_O0_F2.

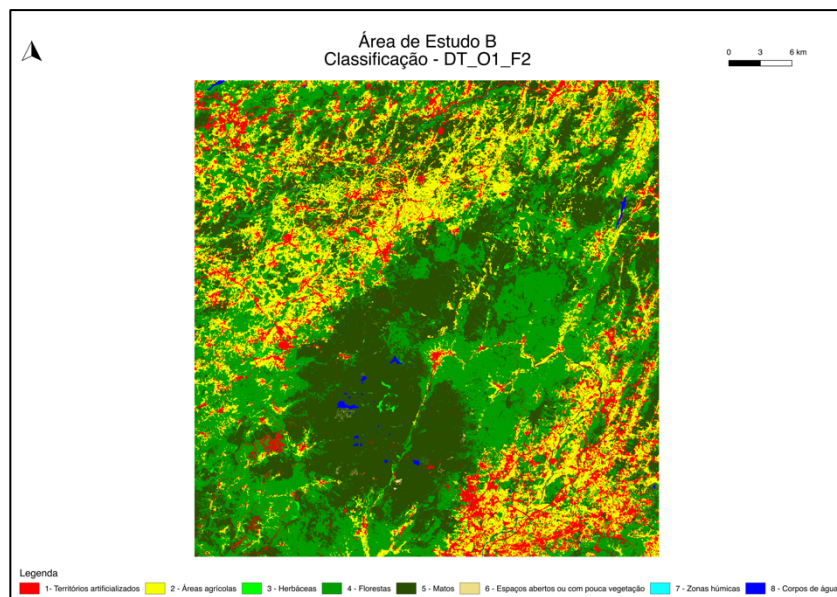


Figura 112: Área de Estudo B – MUCS obtido pela classificação com DT_O1_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,8%	0,2%	0,0%	0,1%	0,2%	0,0%	0,0%	86%
2	4,2%	16,0%	0,0%	1,9%	0,8%	0,0%	0,0%	70%
3	0,3%	1,2%	0,2%	0,2%	0,5%	0,0%	0,0%	8%
4	1,4%	2,4%	0,1%	28,0%	6,7%	0,0%	0,1%	72%
5	1,0%	2,0%	0,0%	4,5%	20,6%	0,0%	0,0%	73%
6	0,0%	0,0%	0,0%	0,1%	3,8%	0,0%	0,0%	1%
8	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,2%	64%
Ex. Utilizador	28%	73%	61%	80%	63%	95%	55%	68%

Tabela 122: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_O1_F2.

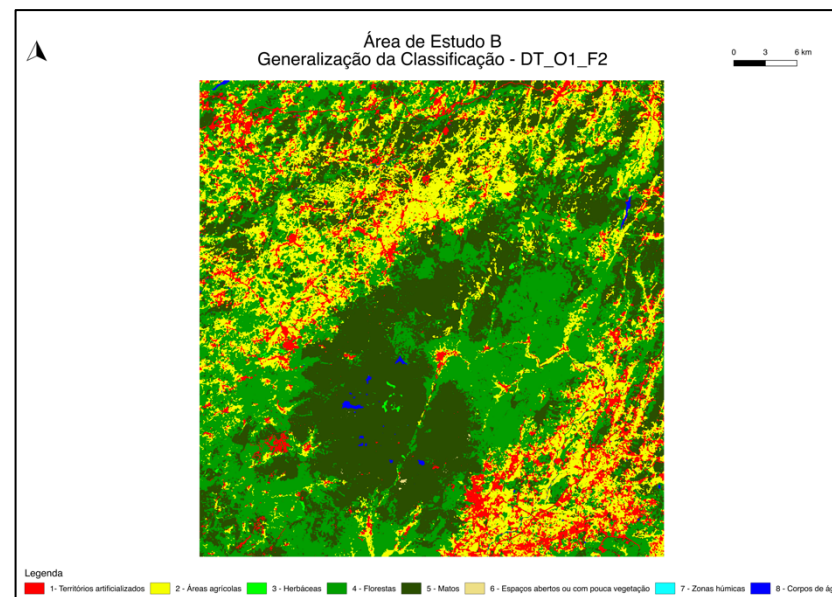


Figura 113: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_O1_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,9%	0,2%	0,0%	0,1%	0,1%	0,0%	0,0%	88%
2	3,8%	16,8%	0,0%	1,8%	0,6%	0,0%	0,0%	73%
3	0,3%	1,3%	0,1%	0,2%	0,6%	0,0%	0,0%	4%
4	1,2%	2,5%	0,1%	28,9%	6,1%	0,0%	0,0%	75%
5	0,7%	2,1%	0,0%	4,1%	21,3%	0,0%	0,0%	76%
6	0,0%	0,0%	0,0%	0,0%	3,9%	0,0%	0,0%	0%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	59%
Ex. Utilizador	32%	73%	65%	82%	65%	99%	77%	70%

Tabela 123: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_O1_F2.

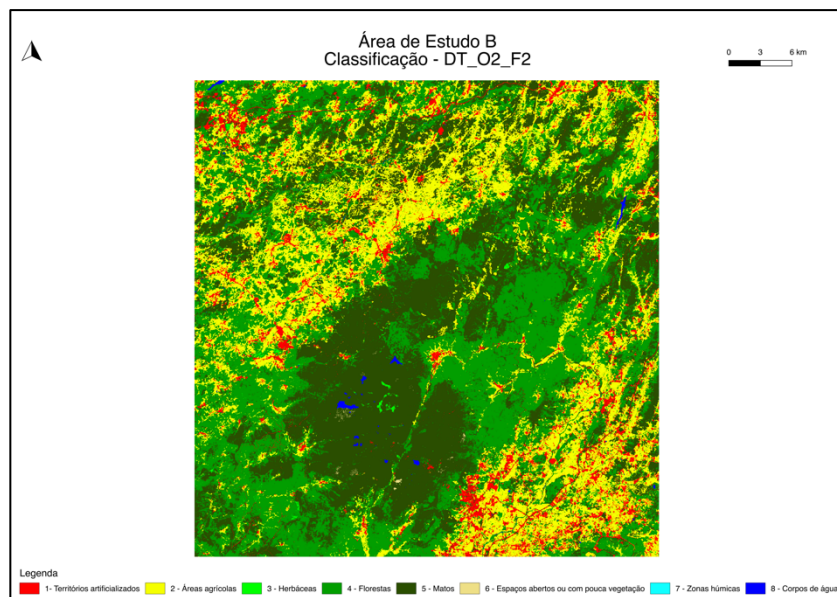


Figura 114: Área de Estudo B – MUCS obtido pela classificação com DT_O2_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,5%	0,4%	0,0%	0,1%	0,3%	0,0%	0,0%	76%
2	2,1%	17,8%	0,0%	2,0%	1,0%	0,0%	0,0%	77%
3	0,2%	1,3%	0,2%	0,2%	0,6%	0,0%	0,0%	8%
4	0,6%	2,6%	0,1%	28,4%	7,1%	0,0%	0,0%	73%
5	0,4%	2,2%	0,0%	4,6%	21,0%	0,0%	0,0%	75%
6	0,0%	0,0%	0,0%	0,1%	3,8%	0,0%	0,0%	1%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	56%
Ex. Utilizador	43%	73%	63%	80%	62%	97%	81%	70%

Tabela 124: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_O2_F2.

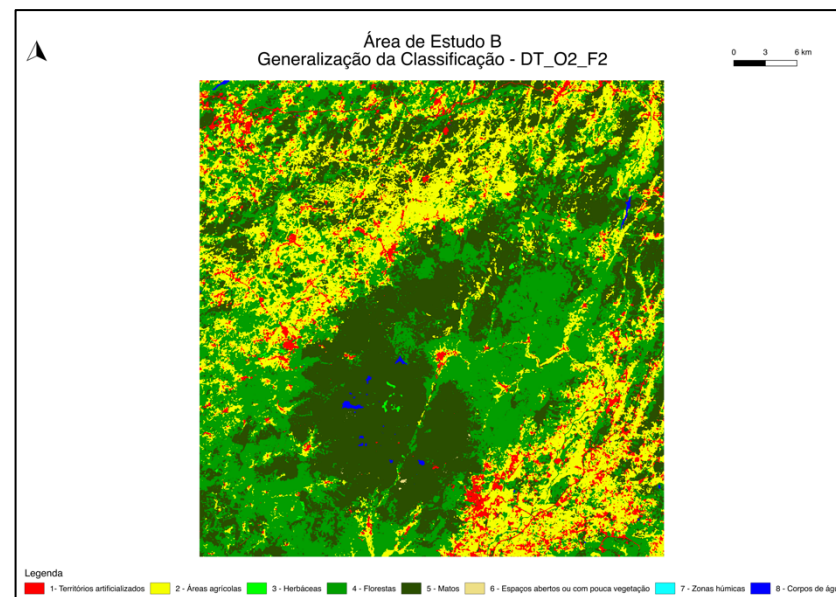


Figura 115: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_O2_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,5%	0,4%	0,0%	0,1%	0,2%	0,0%	0,0%	77%
2	1,8%	18,6%	0,0%	1,9%	0,8%	0,0%	0,0%	81%
3	0,1%	1,4%	0,1%	0,2%	0,6%	0,0%	0,0%	4%
4	0,4%	2,8%	0,0%	29,1%	6,4%	0,0%	0,0%	75%
5	0,2%	2,3%	0,0%	4,1%	21,6%	0,0%	0,0%	77%
6	0,0%	0,0%	0,0%	0,0%	3,9%	0,0%	0,0%	0%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	52%
Ex. Utilizador	50%	73%	67%	82%	64%	99%	86%	72%

Tabela 125: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_O2_F2.

Anexo J – Classificações e respectivas generalizações a partir dos dados de treino obtidos em simultâneo pela COS 2015 e pelo OSM

Neste anexo apresentam-se os MUCS obtidos com os dados de treino extraídos em simultâneo da COS 2015 e do OSM e as respetivas matrizes de confusão.

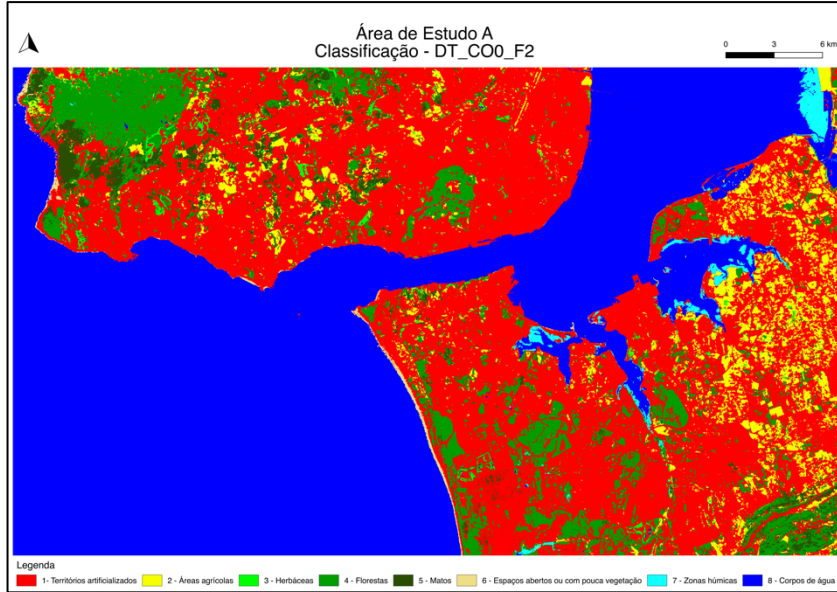


Figura 116: Área de Estudo A – MUCS obtido pela classificação com DT_CO0_F2.

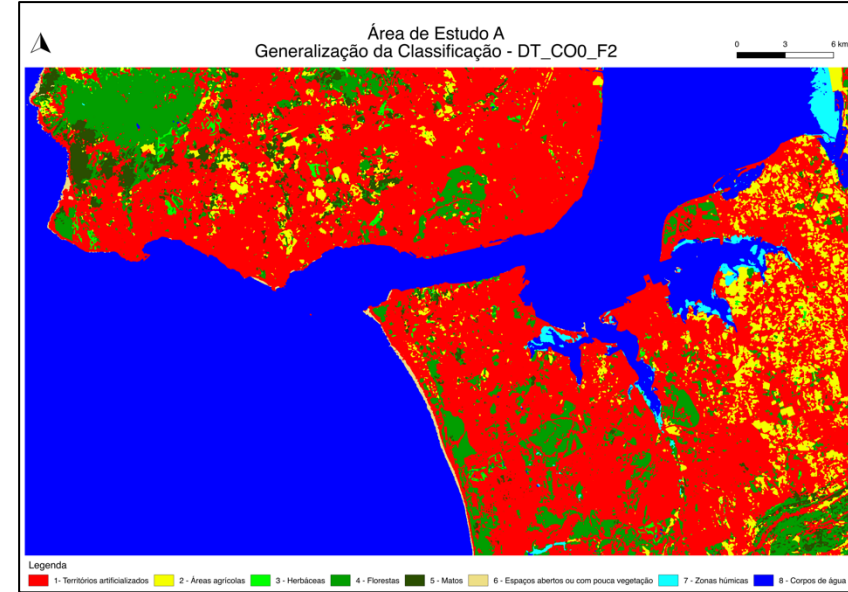


Figura 117: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_CO0_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	23,7%	0,2%	0,1%	0,2%	0,1%	0,0%	0,0%	0,1%	98%
2	4,7%	3,6%	0,3%	0,2%	0,3%	0,0%	0,0%	0,0%	40%
3	1,7%	0,6%	0,4%	0,1%	0,2%	0,0%	0,0%	0,0%	13%
4	5,9%	0,2%	0,1%	7,0%	0,6%	0,0%	0,0%	0,0%	51%
5	1,8%	0,3%	0,1%	0,7%	1,3%	0,0%	0,0%	0,0%	30%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,0%	40%
7	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	73%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,1%	0,2%	43,8%	99%
Ex. Utilizador	62%	73%	43%	85%	52%	54%	68%	99%	81%

Tabela 126: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_CO0_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	23,9%	0,2%	0,0%	0,1%	0,0%	0,0%	0,0%	0,1%	98%
2	5,0%	3,6%	0,1%	0,2%	0,3%	0,0%	0,0%	0,0%	39%
3	1,9%	0,6%	0,3%	0,1%	0,2%	0,0%	0,0%	0,0%	11%
4	6,2%	0,2%	0,0%	7,0%	0,4%	0,0%	0,0%	0,0%	51%
5	2,0%	0,3%	0,1%	0,6%	1,3%	0,0%	0,0%	0,0%	30%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	37%
7	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	74%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,1%	0,2%	43,8%	99%
Ex. Utilizador	61%	75%	54%	86%	59%	57%	72%	99%	81%

Tabela 127: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_CO0_F2.

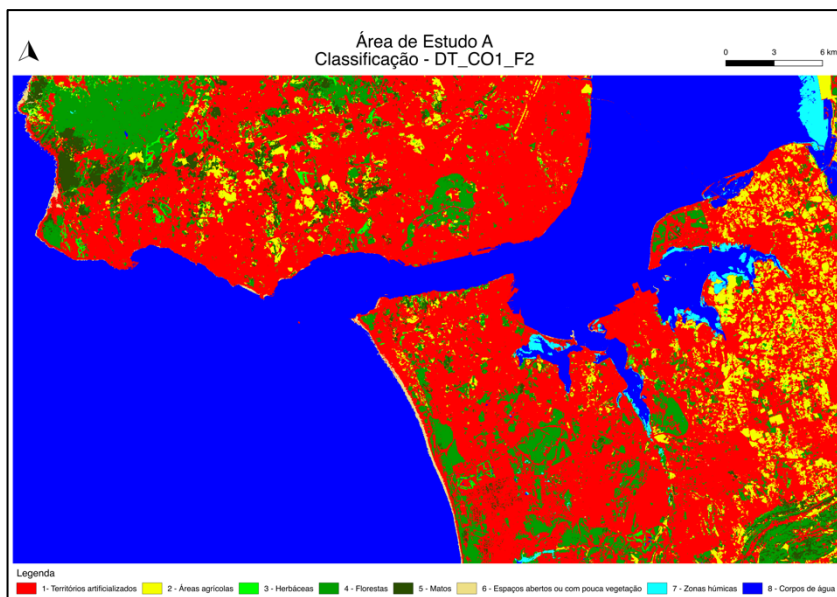


Figura 118: Área de Estudo A – MUCS obtido pela classificação com DT_CO1_F2.

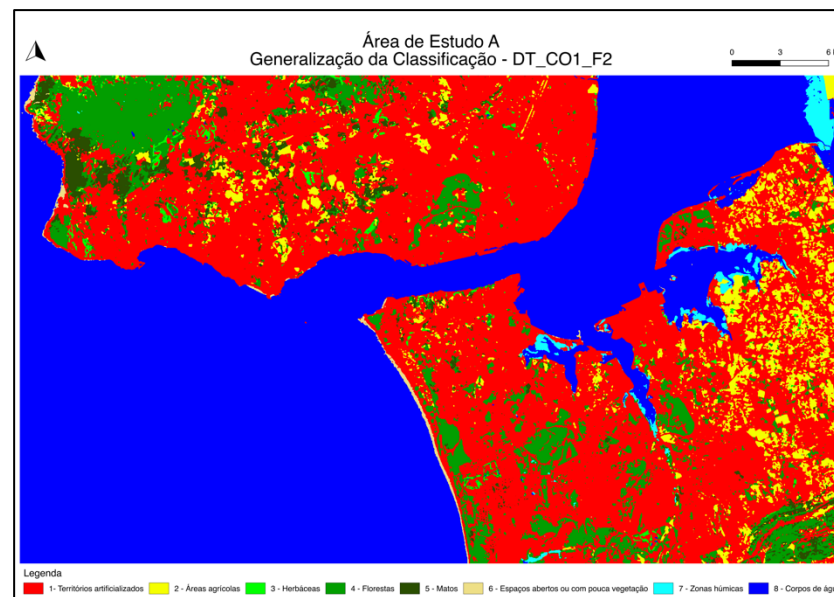


Figura 119: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_CO1_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	23,7%	0,2%	0,1%	0,2%	0,1%	0,0%	0,0%	0,1%	98%
2	4,7%	3,6%	0,3%	0,2%	0,3%	0,0%	0,0%	0,0%	40%
3	1,7%	0,6%	0,4%	0,1%	0,2%	0,0%	0,0%	0,0%	13%
4	5,9%	0,2%	0,1%	7,0%	0,6%	0,0%	0,0%	0,0%	51%
5	1,8%	0,3%	0,1%	0,7%	1,3%	0,0%	0,0%	0,0%	30%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,0%	40%
7	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	73%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,1%	0,2%	43,7%	99%
Ex. Utilizador	62%	74%	43%	85%	52%	53%	68%	99%	81%

Tabela 128: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_CO1_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	23,9%	0,1%	0,0%	0,1%	0,0%	0,0%	0,0%	0,1%	98%
2	5,0%	3,5%	0,1%	0,2%	0,3%	0,0%	0,0%	0,0%	39%
3	1,9%	0,6%	0,4%	0,1%	0,2%	0,0%	0,0%	0,0%	11%
4	6,1%	0,2%	0,0%	7,0%	0,4%	0,0%	0,0%	0,0%	51%
5	2,0%	0,3%	0,1%	0,6%	1,3%	0,0%	0,0%	0,0%	30%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	37%
7	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	74%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,1%	0,2%	43,8%	99%
Ex. Utilizador	61%	76%	53%	86%	59%	55%	71%	99%	81%

Tabela 129: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_CO1_F2.

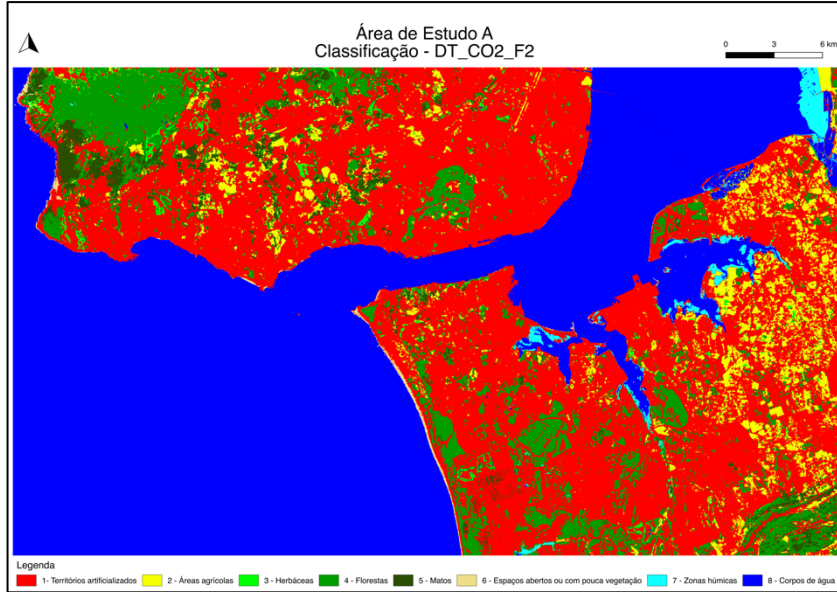


Figura 120: Área de Estudo A – MUCS obtido pela classificação com DT_CO2_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	23,7%	0,2%	0,1%	0,2%	0,1%	0,0%	0,0%	0,1%	97%
2	4,7%	3,5%	0,4%	0,2%	0,4%	0,0%	0,0%	0,0%	38%
3	1,7%	0,6%	0,4%	0,2%	0,2%	0,0%	0,0%	0,0%	14%
4	5,7%	0,2%	0,1%	7,1%	0,6%	0,0%	0,0%	0,0%	52%
5	1,8%	0,3%	0,2%	0,7%	1,4%	0,0%	0,0%	0,0%	32%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	37%
7	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	72%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,1%	0,2%	43,7%	99%
Ex. Utilizador	62%	74%	38%	84%	51%	59%	66%	99%	80%

Tabela 130: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_CO2_F2.

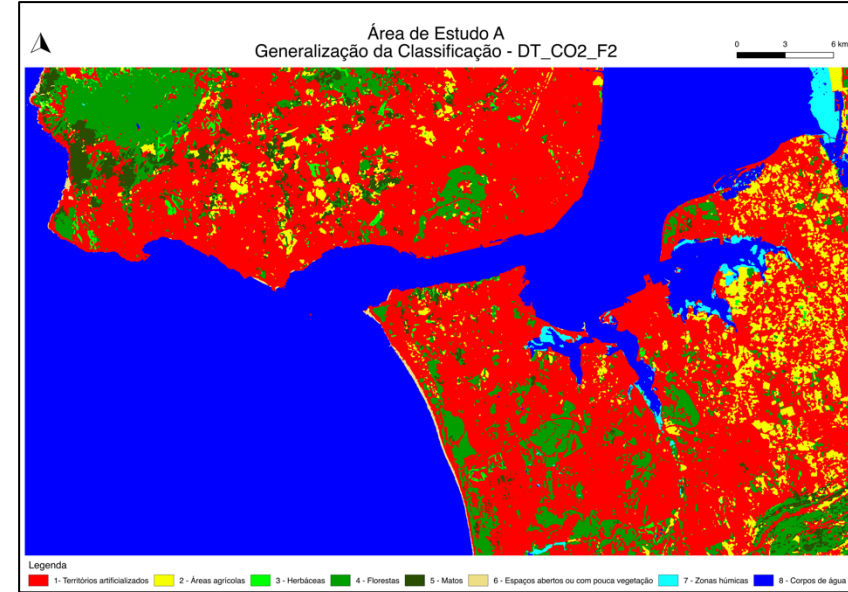


Figura 121: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_CO2_F2.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	23,8%	0,1%	0,1%	0,2%	0,0%	0,0%	0,0%	0,1%	98%
2	5,0%	3,4%	0,2%	0,2%	0,3%	0,0%	0,0%	0,0%	37%
3	1,8%	0,5%	0,4%	0,1%	0,2%	0,0%	0,0%	0,0%	12%
4	6,0%	0,1%	0,0%	7,1%	0,4%	0,0%	0,0%	0,0%	52%
5	1,9%	0,2%	0,1%	0,7%	1,3%	0,0%	0,0%	0,0%	31%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	34%
7	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,6%	0,1%	73%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,1%	0,2%	43,8%	99%
Ex. Utilizador	61%	76%	47%	85%	58%	63%	70%	99%	81%

Tabela 131: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_CO2_F2.

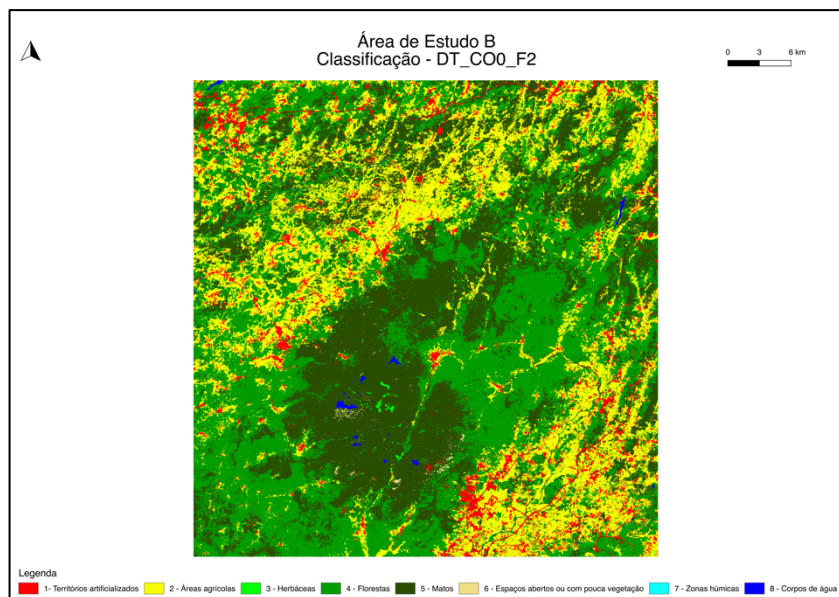


Figura 122: Área de Estudo B – MUCS obtido pela classificação com DT_CO0_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,6%	0,4%	0,0%	0,1%	0,2%	0,0%	0,0%	78%
2	1,8%	17,7%	0,0%	2,5%	1,0%	0,0%	0,0%	77%
3	0,2%	1,3%	0,2%	0,3%	0,5%	0,0%	0,0%	7%
4	0,5%	2,4%	0,0%	29,8%	6,1%	0,0%	0,0%	77%
5	0,3%	2,1%	0,0%	5,6%	20,1%	0,0%	0,0%	71%
6	0,0%	0,0%	0,0%	0,3%	3,5%	0,1%	0,0%	3%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	60%
Ex. Utilizador	48%	74%	88%	77%	64%	96%	77%	71%

Tabela 132: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_CO0_F2.

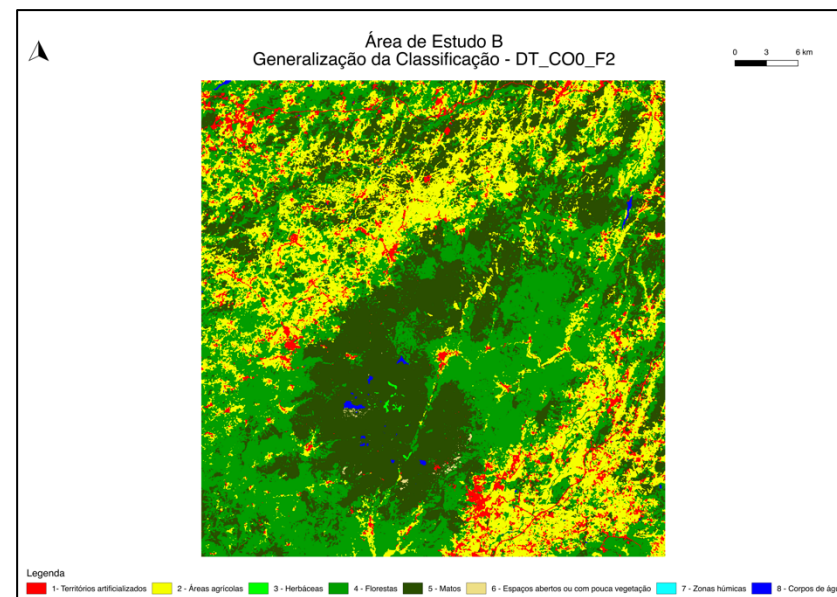


Figura 123: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_CO0_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,6%	0,4%	0,0%	0,1%	0,1%	0,0%	0,0%	79%
2	1,5%	18,4%	0,0%	2,4%	0,8%	0,0%	0,0%	80%
3	0,1%	1,4%	0,1%	0,3%	0,6%	0,0%	0,0%	4%
4	0,4%	2,5%	0,0%	30,6%	5,2%	0,0%	0,0%	79%
5	0,3%	2,2%	0,0%	5,1%	20,6%	0,0%	0,0%	73%
6	0,0%	0,0%	0,0%	0,3%	3,6%	0,1%	0,0%	2%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	55%
Ex. Utilizador	52%	74%	94%	79%	67%	99%	83%	73%

Tabela 133: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_CO0_F2.

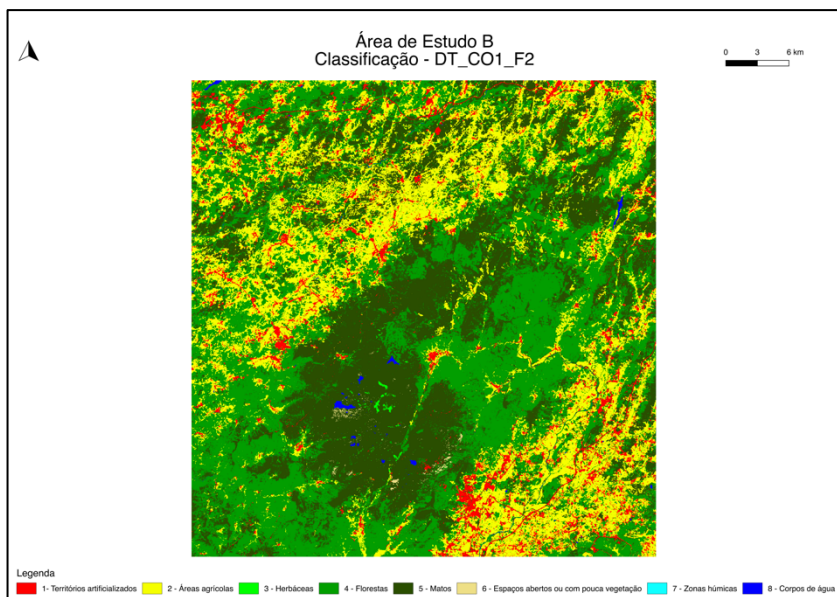


Figura 124: Área de Estudo B – MUCS obtido pela classificação com DT_CO1_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,6%	0,4%	0,0%	0,1%	0,2%	0,0%	0,0%	79%
2	1,8%	17,7%	0,0%	2,5%	1,0%	0,0%	0,0%	77%
3	0,2%	1,3%	0,2%	0,3%	0,5%	0,0%	0,0%	7%
4	0,5%	2,4%	0,0%	29,9%	6,1%	0,0%	0,0%	77%
5	0,4%	2,2%	0,0%	5,6%	20,0%	0,0%	0,0%	71%
6	0,0%	0,0%	0,0%	0,3%	3,5%	0,1%	0,0%	3%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	60%
Ex. Utilizador	47%	74%	88%	77%	64%	96%	77%	71%

Tabela 134: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_CO1_F2.

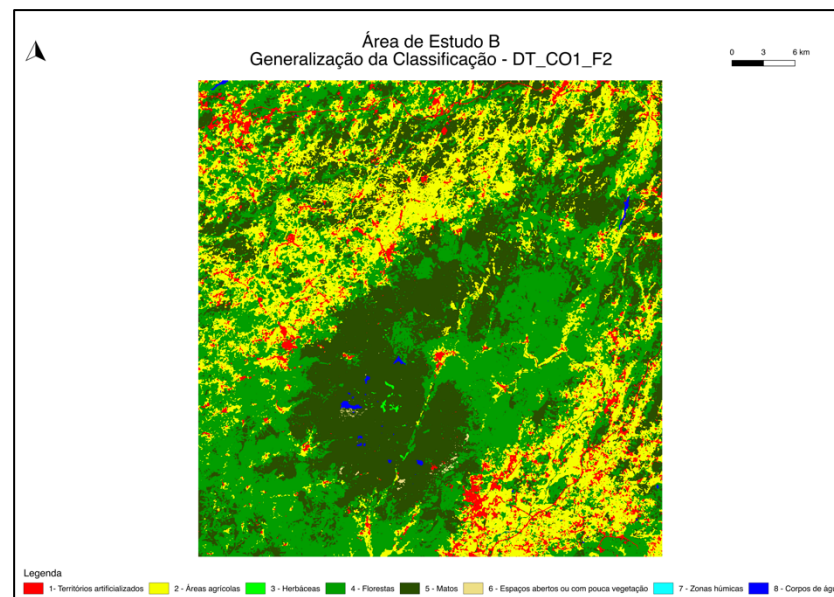


Figura 125: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_CO1_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,6%	0,4%	0,0%	0,1%	0,1%	0,0%	0,0%	80%
2	1,5%	18,4%	0,0%	2,4%	0,7%	0,0%	0,0%	80%
3	0,1%	1,4%	0,1%	0,3%	0,6%	0,0%	0,0%	4%
4	0,4%	2,5%	0,0%	30,7%	5,2%	0,0%	0,0%	79%
5	0,3%	2,2%	0,0%	5,1%	20,6%	0,0%	0,0%	73%
6	0,0%	0,0%	0,0%	0,3%	3,6%	0,1%	0,0%	2%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	55%
Ex. Utilizador	52%	74%	94%	79%	67%	99%	83%	73%

Tabela 135: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_CO1_F2.

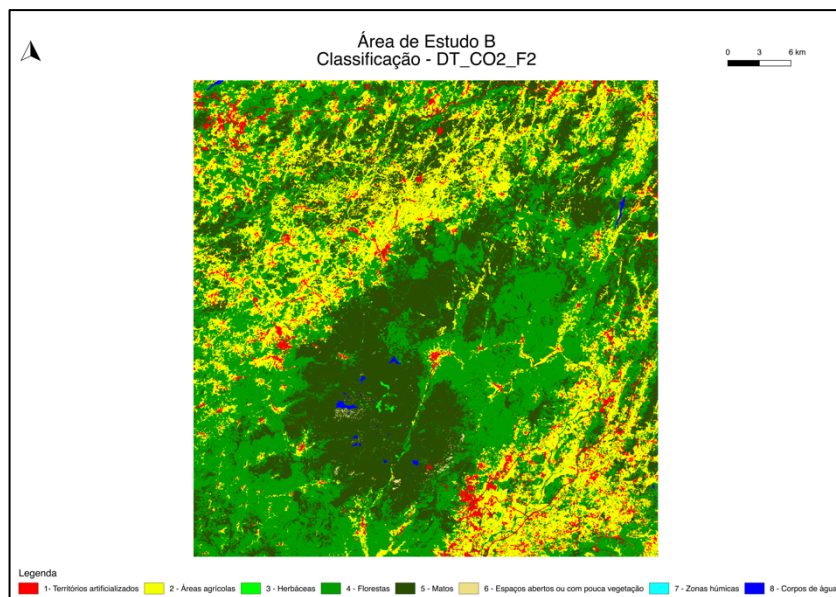


Figura 126: Área de Estudo B – MUCS obtido pela classificação com DT_CO2_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,3%	0,5%	0,0%	0,1%	0,3%	0,0%	0,0%	71%
2	1,3%	18,1%	0,0%	2,5%	1,1%	0,0%	0,0%	79%
3	0,1%	1,3%	0,1%	0,3%	0,6%	0,0%	0,0%	6%
4	0,3%	2,4%	0,0%	29,8%	6,2%	0,0%	0,0%	77%
5	0,2%	2,2%	0,0%	5,7%	20,1%	0,0%	0,0%	71%
6	0,0%	0,0%	0,0%	0,3%	3,5%	0,1%	0,0%	3%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	55%
Ex. Utilizador	54%	74%	90%	77%	63%	96%	83%	71%

Tabela 136: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_CO2_F2.

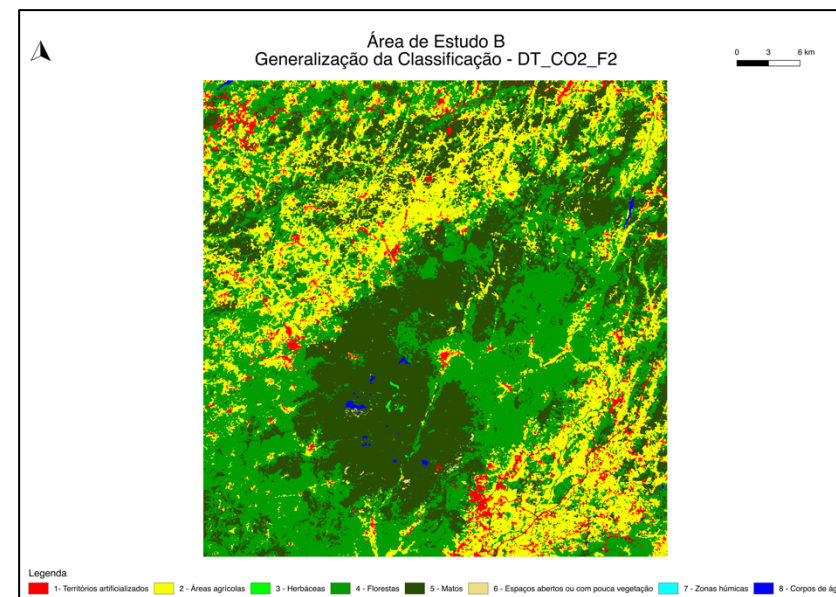


Figura 127: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_CO2_F2.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	2,3%	0,5%	0,0%	0,2%	0,2%	0,0%	0,0%	72%
2	1,1%	18,8%	0,0%	2,4%	0,8%	0,0%	0,0%	81%
3	0,1%	1,4%	0,1%	0,3%	0,6%	0,0%	0,0%	4%
4	0,3%	2,6%	0,0%	30,6%	5,3%	0,0%	0,0%	79%
5	0,1%	2,2%	0,0%	5,2%	20,6%	0,0%	0,0%	73%
6	0,0%	0,0%	0,0%	0,3%	3,6%	0,1%	0,0%	2%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	52%
Ex. Utilizador	59%	74%	94%	79%	66%	99%	86%	73%

Tabela 137: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_CO2_F2.

Anexo K – Classificações e respectivas generalizações a partir dos dados de treino extraídos do UA 2012

Neste anexo apresentam-se os MUCS obtidos com os dados de treino extraídos do UA 2012 e as respectivas matrizes de confusão.

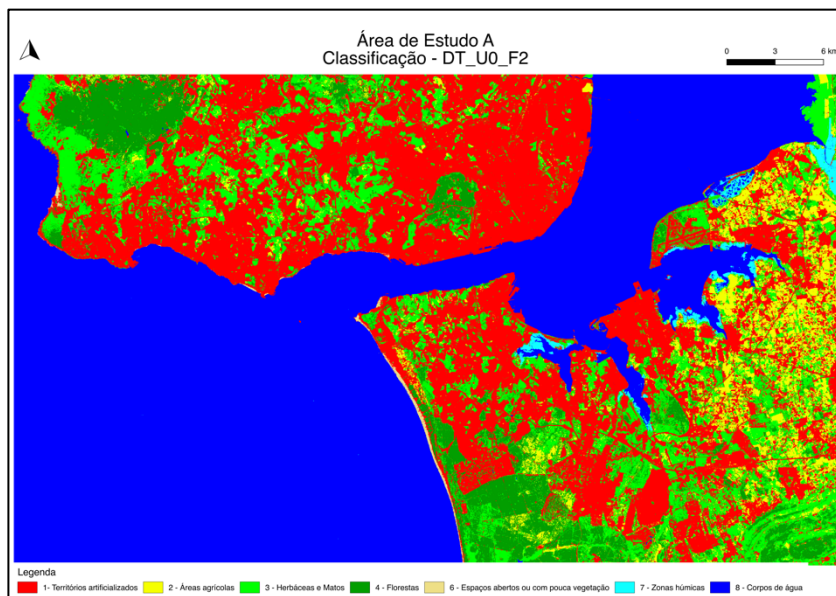


Figura 128: Área de Estudo A – MUCS obtido pela classificação com DT_U0_F2.

Classe	1	2	3	4	6	7	8	Ex. Produtor
1	22,8%	0,3%	1,0%	0,1%	0,0%	0,0%	0,1%	94%
2	2,2%	3,4%	3,5%	0,1%	0,0%	0,0%	0,0%	37%
3	2,2%	0,5%	4,3%	0,4%	0,0%	0,0%	0,0%	58%
4	1,9%	0,7%	5,0%	6,2%	0,0%	0,0%	0,0%	45%
6	0,1%	0,0%	0,0%	0,0%	0,1%	0,0%	0,1%	35%
7	0,0%	0,0%	0,3%	0,0%	0,0%	0,3%	0,1%	35%
8	0,1%	0,0%	0,1%	0,0%	0,0%	0,4%	43,7%	99%
Ex. Utilizador	78%	68%	30%	92%	80%	41%	99%	81%

Tabela 138: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_U0_F2.

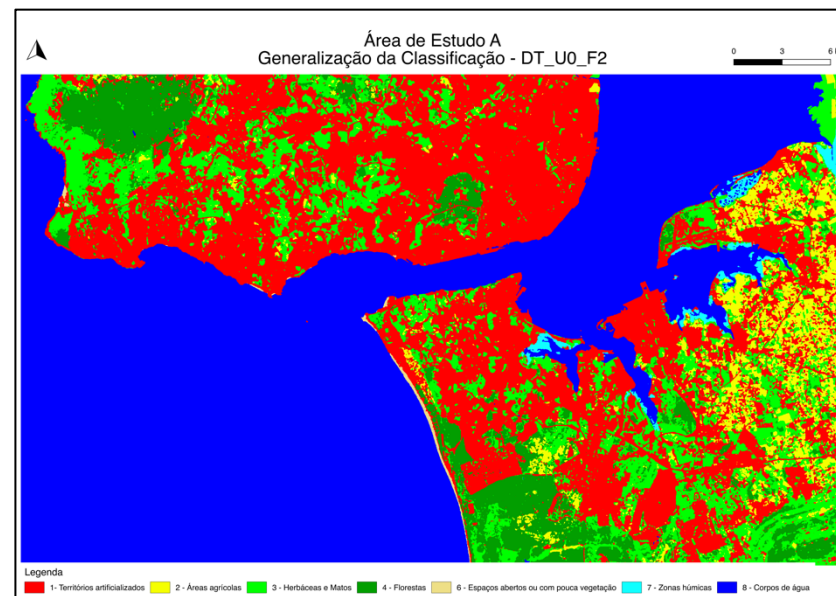


Figura 129: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_U0_F2.

Classe	1	2	3	4	6	7	8	Ex. Produtor
1	23,2%	0,2%	0,8%	0,1%	0,0%	0,0%	0,1%	95%
2	2,3%	3,3%	3,5%	0,0%	0,0%	0,0%	0,0%	36%
3	2,4%	0,4%	4,3%	0,3%	0,0%	0,0%	0,0%	58%
4	2,0%	0,5%	5,1%	6,2%	0,0%	0,0%	0,0%	45%
6	0,2%	0,0%	0,0%	0,0%	0,1%	0,0%	0,1%	34%
7	0,0%	0,0%	0,3%	0,0%	0,0%	0,3%	0,1%	36%
8	0,0%	0,0%	0,0%	0,0%	0,0%	0,4%	43,7%	99%
Ex. Utilizador	77%	75%	31%	93%	85%	42%	99%	81%

Tabela 139: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_U0_F2.

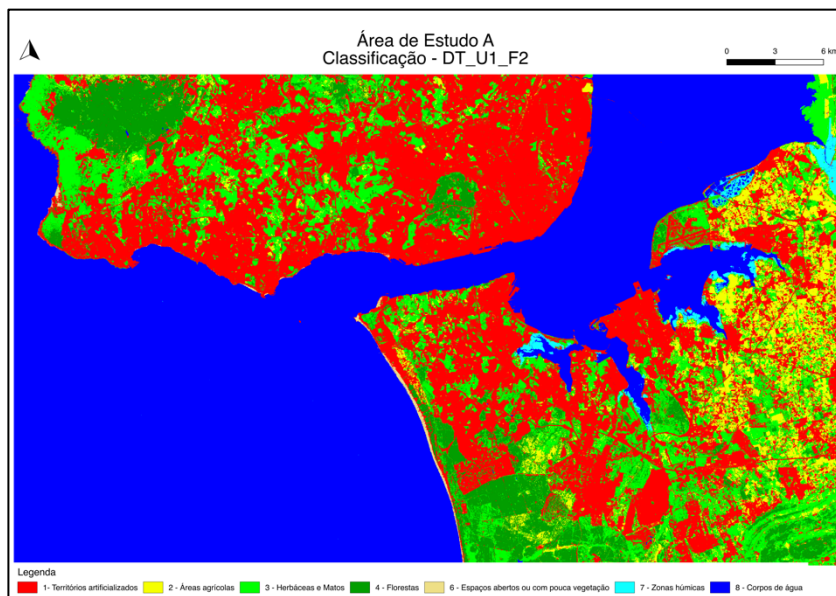


Figura 130: Área de Estudo A – MUCS obtido pela classificação com DT_U1_F2.

Classe	1	2	3	4	6	7	8	Ex. Produtor
1	22,7%	0,3%	1,1%	0,1%	0,0%	0,0%	0,1%	93%
2	2,1%	3,4%	3,5%	0,1%	0,0%	0,0%	0,0%	38%
3	2,1%	0,6%	4,3%	0,4%	0,0%	0,0%	0,0%	58%
4	1,8%	0,7%	5,0%	6,2%	0,0%	0,0%	0,0%	45%
6	0,1%	0,0%	0,0%	0,0%	0,1%	0,0%	0,1%	36%
7	0,0%	0,0%	0,3%	0,0%	0,0%	0,3%	0,1%	36%
8	0,1%	0,0%	0,1%	0,0%	0,0%	0,4%	43,7%	99%
Ex. Utilizador	78%	68%	30%	92%	80%	42%	99%	81%

Tabela 140: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_U1_F2.

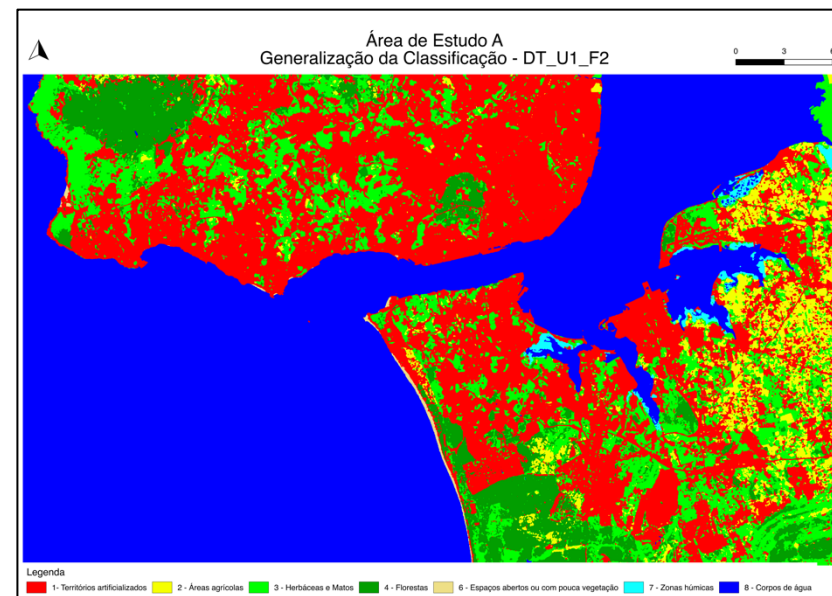


Figura 131: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_U1_F2.

Classe	1	2	3	4	6	7	8	Ex. Produtor
1	23,1%	0,2%	0,8%	0,1%	0,0%	0,0%	0,1%	95%
2	2,2%	3,4%	3,5%	0,0%	0,0%	0,0%	0,0%	37%
3	2,3%	0,4%	4,4%	0,3%	0,0%	0,0%	0,0%	59%
4	1,9%	0,5%	5,1%	6,2%	0,0%	0,0%	0,0%	45%
6	0,2%	0,0%	0,0%	0,0%	0,1%	0,0%	0,1%	34%
7	0,0%	0,0%	0,3%	0,0%	0,0%	0,3%	0,1%	36%
8	0,0%	0,0%	0,0%	0,0%	0,0%	0,4%	43,7%	99%
Ex. Utilizador	78%	75%	31%	93%	85%	42%	99%	81%

Tabela 141: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_U1_F2.

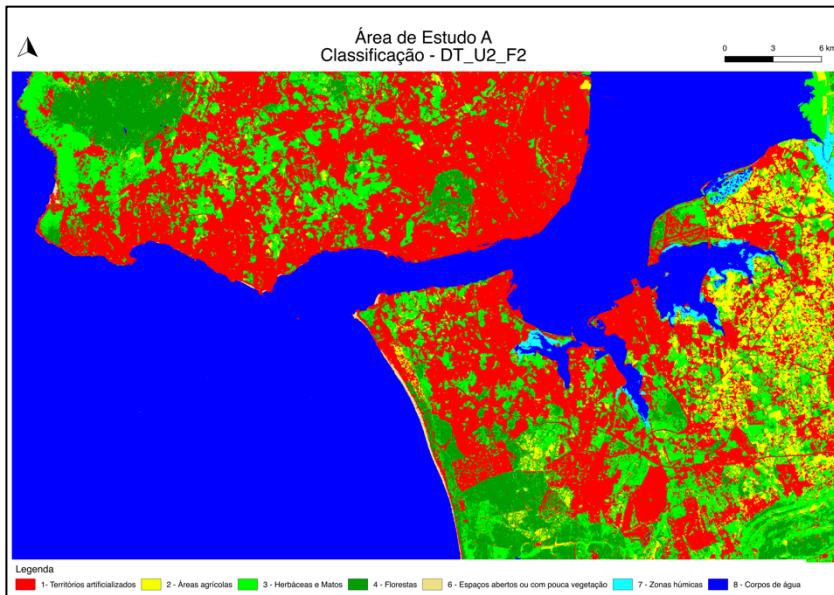


Figura 132: Área de Estudo A – MUCS obtido pela classificação com DT_U2_F2.

Classe	1	2	3	4	6	7	8	Ex. Produtor
1	22,8%	0,3%	1,1%	0,1%	0,0%	0,0%	0,1%	94%
2	2,1%	3,4%	3,6%	0,1%	0,0%	0,0%	0,0%	37%
3	2,2%	0,5%	4,3%	0,4%	0,0%	0,0%	0,0%	58%
4	1,8%	0,7%	5,1%	6,2%	0,0%	0,0%	0,0%	45%
6	0,2%	0,0%	0,0%	0,0%	0,1%	0,0%	0,1%	35%
7	0,0%	0,0%	0,3%	0,0%	0,0%	0,3%	0,1%	36%
8	0,1%	0,0%	0,1%	0,0%	0,0%	0,4%	43,7%	99%
Ex. Utilizador	78%	69%	30%	92%	81%	41%	99%	81%

Tabela 142: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_U2_F2.

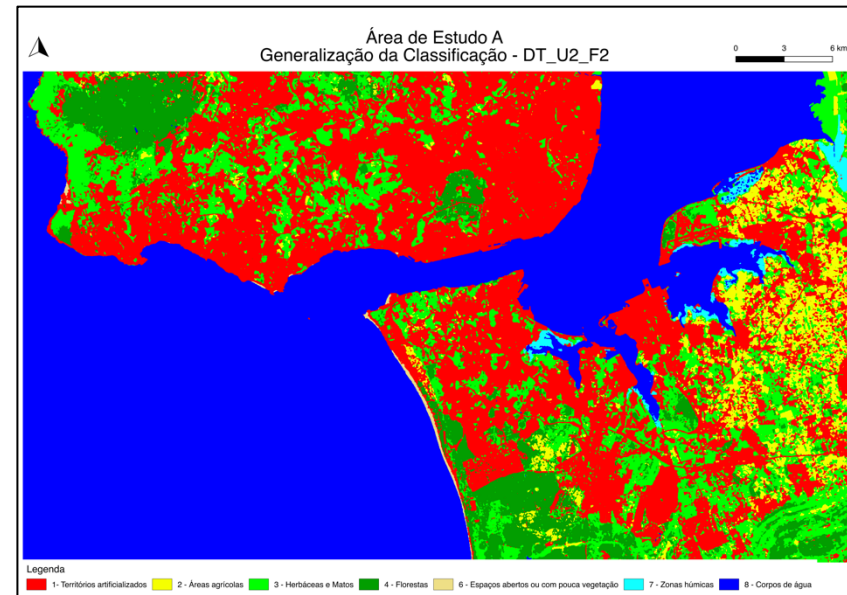


Figura 133: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_U2_F2.

Classe	1	2	3	4	6	7	8	Ex. Produtor
1	23,1%	0,2%	0,8%	0,1%	0,0%	0,0%	0,1%	95%
2	2,2%	3,4%	3,5%	0,0%	0,0%	0,0%	0,0%	37%
3	2,4%	0,4%	4,4%	0,3%	0,0%	0,0%	0,0%	59%
4	1,9%	0,5%	5,2%	6,2%	0,0%	0,0%	0,0%	45%
6	0,2%	0,0%	0,0%	0,0%	0,1%	0,0%	0,1%	33%
7	0,0%	0,0%	0,3%	0,0%	0,0%	0,3%	0,1%	37%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,4%	43,7%	99%
Ex. Utilizador	77%	75%	31%	93%	85%	42%	99%	81%

Tabela 143: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_U2_F2.

Anexo L – Classificações e Respetivas Generalizações a Partir dos Dados de Treino Extraídos da COS 2015 Utilizando os Vizinhos como *Features*

Neste anexo apresentam-se os MUCS obtidos com os dados de treino extraídos da COS 2015 usando os vizinhos como *features* e as respetivas matrizes de confusão.

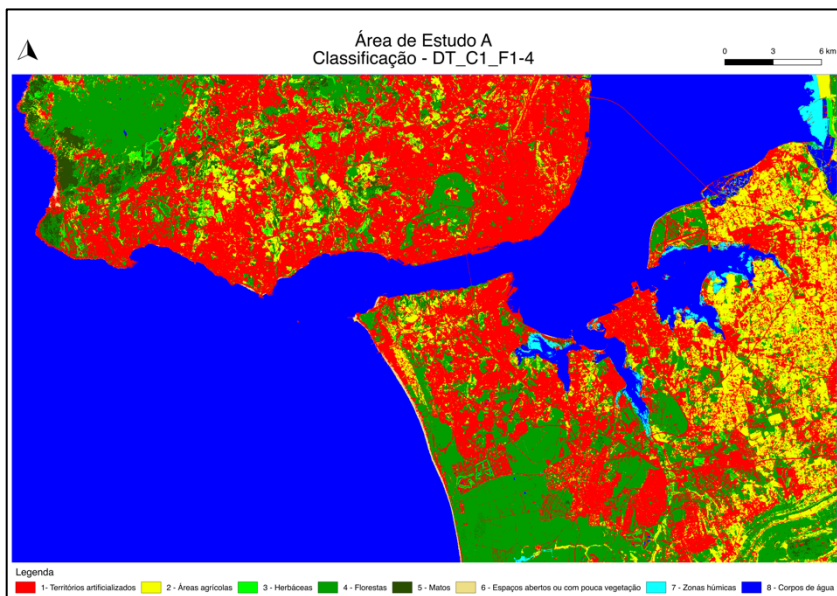


Figura 134: Área de Estudo A – MUCS obtido pela classificação com DT_C1_F1-4.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	22,1%	0,9%	0,2%	1,0%	0,1%	0,0%	0,0%	0,1%	91%
2	1,9%	6,0%	0,4%	0,7%	0,2%	0,0%	0,0%	0,0%	66%
3	0,8%	1,0%	0,9%	0,4%	0,1%	0,0%	0,0%	0,0%	28%
4	1,4%	0,5%	0,1%	11,5%	0,2%	0,0%	0,0%	0,0%	83%
5	0,8%	0,7%	0,3%	1,1%	1,4%	0,0%	0,0%	0,0%	32%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,2%	0,0%	0,0%	45%
7	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,6%	0,1%	70%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,1%	44,0%	100%
Ex. Utilizador	81%	66%	47%	78%	71%	87%	89%	99%	86%

Tabela 144: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1_F1-4.

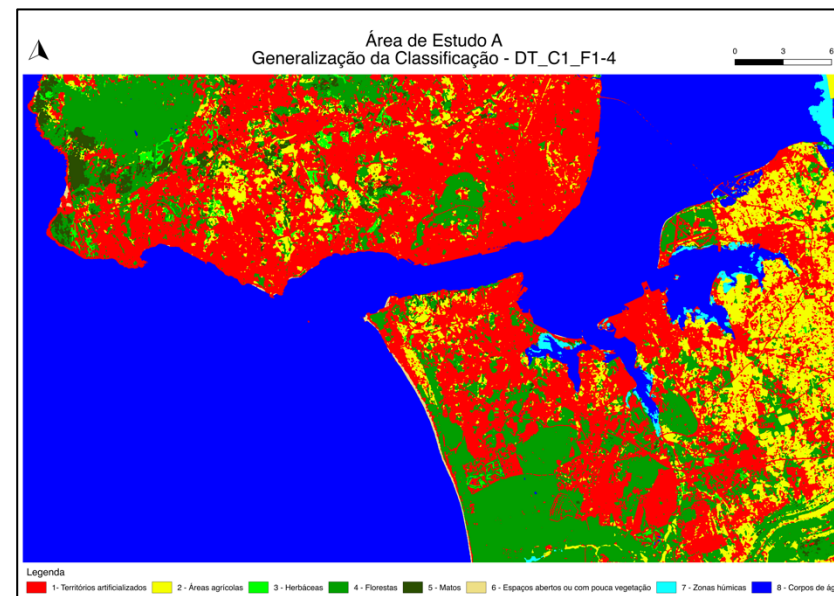


Figura 135: Área de Estudo A – MUCS obtido pela generalização da classificação com DT_C1_F1-4.

Classe	1	2	3	4	5	6	7	8	Ex. Produtor
1	22,9%	0,5%	0,0%	0,8%	0,0%	0,0%	0,0%	0,1%	94%
2	1,9%	6,4%	0,1%	0,6%	0,1%	0,0%	0,0%	0,0%	70%
3	1,0%	1,0%	0,7%	0,4%	0,1%	0,0%	0,0%	0,0%	21%
4	1,4%	0,4%	0,0%	11,8%	0,1%	0,0%	0,0%	0,0%	86%
5	1,2%	0,6%	0,2%	1,2%	1,2%	0,0%	0,0%	0,0%	28%
6	0,2%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	38%
7	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,6%	0,1%	72%
8	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	44,0%	100%
Ex. Utilizador	80%	72%	65%	79%	82%	94%	92%	99%	88%

Tabela 145: Área de Estudo A – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1_F1-4.

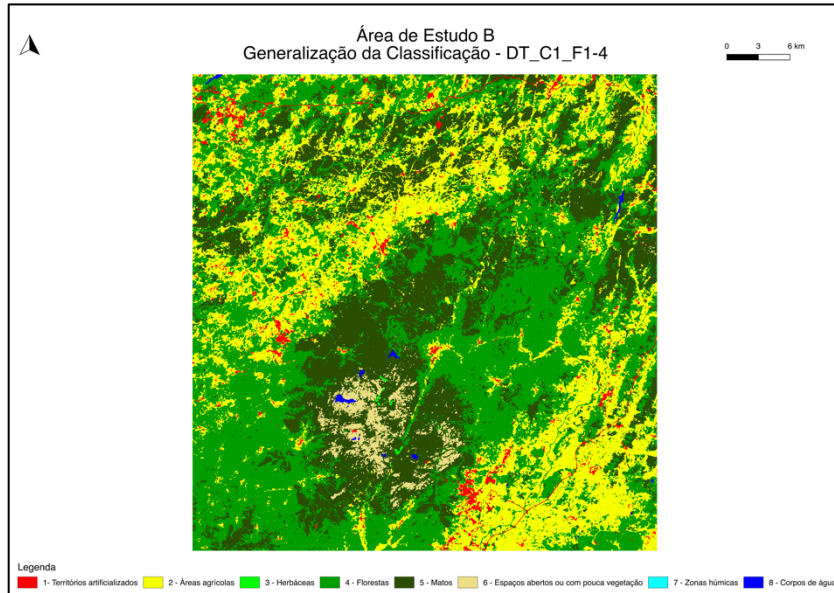


Figura 136: Área de Estudo B – MUCS obtido pela classificação com DT_C1_F1-4.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	1,9%	0,9%	0,0%	0,3%	0,2%	0,0%	0,0%	57%
2	0,3%	18,8%	0,2%	2,6%	1,1%	0,0%	0,0%	82%
3	0,0%	1,2%	0,3%	0,3%	0,6%	0,0%	0,0%	10%
4	0,1%	2,3%	0,0%	31,3%	5,0%	0,0%	0,0%	81%
5	0,1%	1,7%	0,1%	6,7%	19,1%	0,4%	0,0%	68%
6	0,0%	0,0%	0,0%	0,3%	2,0%	1,7%	0,0%	43%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	56%
Ex. Utilizador	76%	75%	47%	75%	68%	77%	62%	73%

Tabela 146: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela classificação com DT_C1_F1-4.

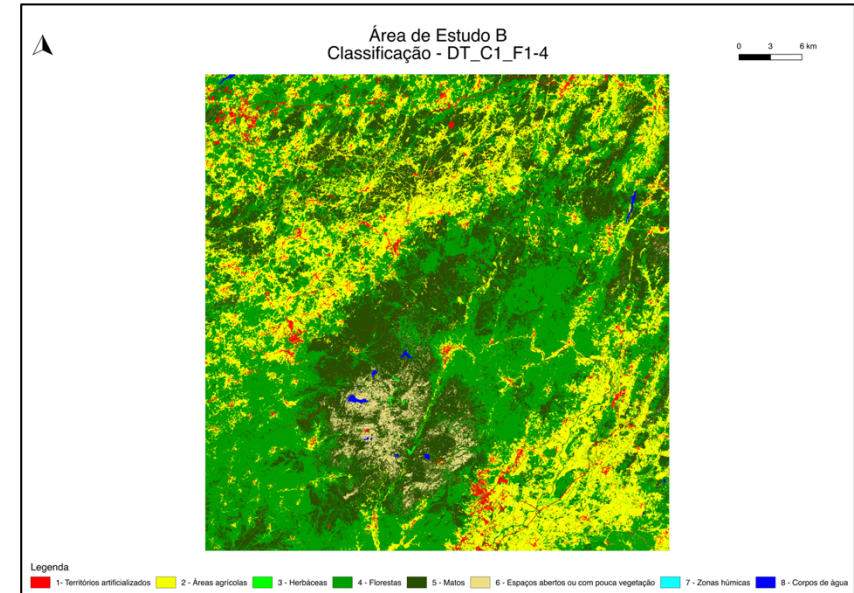


Figura 137: Área de Estudo B – MUCS obtido pela generalização da classificação com DT_C1_F1-4.

Classe	1	2	3	4	5	6	8	Ex. Produtor
1	1,8%	1,1%	0,0%	0,3%	0,1%	0,0%	0,0%	55%
2	0,1%	19,7%	0,0%	2,5%	0,7%	0,0%	0,0%	85%
3	0,0%	1,4%	0,1%	0,3%	0,6%	0,0%	0,0%	3%
4	0,1%	2,4%	0,0%	32,2%	4,1%	0,0%	0,0%	83%
5	0,0%	1,7%	0,0%	6,2%	19,9%	0,2%	0,0%	71%
6	0,0%	0,0%	0,0%	0,2%	2,2%	1,6%	0,0%	40%
8	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,2%	50%
Ex. Utilizador	86%	75%	89%	77%	72%	85%	93%	75%

Tabela 147: Área de Estudo B – Matriz de confusão do processo de validação do MUCS obtido pela generalização da classificação com DT_C1_F1-4.