

Entregáveis E.2.1. e E.2.2. – Contrato Nº 24IN10150012

Metadados – Base de Dados de Referência para validação de mapas de perdas de vegetação

António Tomás Oruguela Sequeira

setembro de 2024

Nota prévia

Este trabalho foi realizado no âmbito do contrato Nº 24IN10150012, enquadrado no contrato de cooperação Nº 3044, celebrado entre a Direção-Geral do Território (DGT) e o Instituto Superior de Agronomia (ISA).

O presente documento corresponde aos Entregáveis E.2.1. e E.2.2. que integram a Tarefa 2, que tem a seguinte descrição: “Apoio à construção de uma base de dados de referência (BDR) para calibração e validação espacial e temporal das metodologias a desenvolver. Recolha e organização de dados de acordo com um delineamento para Portugal Continental.”

Este documento descreve a base de dados de referência (BDR) produzida, que foi baseada em dados fornecidos pela The Navigator Company e tem como finalidade a validação de mapas de perdas de vegetação.

Índice

Enquadramento	4
Conteúdo	4
Área de estudo	5
Dados base	5
Componente espacial	6
Tabelas auxiliares	7
Metodologia	9
Problema inicial e abordagem seguida	9
Divisão dos sub-talhões	9
Atribuição de uma data de corte a cada sub-talhão	9
Tabela de atributos	10
Limitações da BDR	11

Enquadramento

A BDR foi desenvolvida a partir dos dados fornecidos pela The Navigator Company com o objetivo de validar mapas de perdas de vegetação.

Os dados originais, que foram fornecidos à DGT, dizem respeito aos povoamentos florestais da The Navigator Company e contêm informações sobre a ocupação do solo e as operações de exploração realizadas nos povoamentos, incluindo cortes rasos de floresta. Como os dados não foram desenvolvidos com o intuito de validar ou treinar algoritmos de deteção de perdas de vegetação, foram necessários vários processos para fazer a sua adaptação aos objetivos pretendidos.

No presente documento estão descritos os dados originais fornecidos, a base de dados de referência final e os métodos aplicados para a sua obtenção.

Conteúdo

Uma base de dados que compreende 13 418 polígonos em formato vetorial Geopackage que está projetada no sistema de coordenadas CRS: EPSG: 3763 – ETRS89 / PTM06 e o ficheiro de metadados em formato PDF (este documento) a descrever os dados e a sua metodologia de produção.

Os ficheiros disponibilizados são os seguintes:

- E_2_1_BDR_AntonioSequeira.gpkg
- E.2.2._MetadadosBDR_AntónioSequeira.pdf (este documento)

Área de estudo

A área coberta pelos dados corresponde ao território de Portugal Continental (Figura 1).

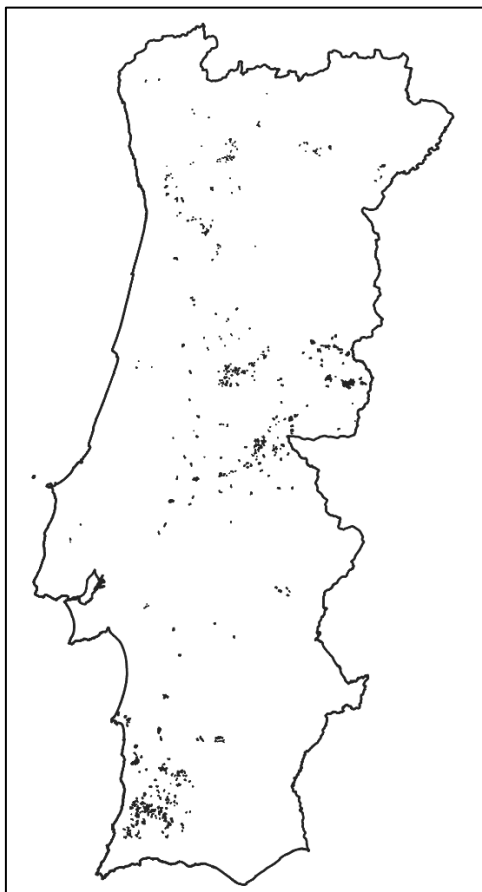


Figura 1 - Áreas cobertas pela BDR

Dados base

Os dados base utilizados para gerar a BDR dizem respeito às parcelas fornecidas pela The Navigator Company e apresentam as especificações que se encontram na Tabela 1.

Especificações	
Formato	Vetorial – polígonos (GPKG)
Nº de features/talhões	1 583
Área total	28 544 ha
Ocupações/Espécies	Eucalipto (91%); Pinheiro bravo (4%); Pinheiro manso (2%); outros pinheiros (2%)
Nº de campos na tabela de atributos	12
Nome dos ficheiros	NVG_2015-2023_Próprios_clean, Exploracao_NVG_2015-2023_Proprios_clean, Silvicultura_NVG_2015-2023_Proprios_clean

Tabela 1 - Especificações dos dados base

Os dados incluem a delimitação de povoamentos florestais em Portugal Continental e são acompanhados por três tabelas que complementam a informação espacial (a tabela de atributos que acompanha a camada com dimensão espacial e outras duas sem componente espacial).

Componente espacial

Camada “NVG_2015-2023_Próprios_clean”: contem as *features* que delimitam as áreas florestais. Para cada *feature*, correspondente a um talhão (unidade mínima de gestão da The Navigator Company), existem as informações que se encontram na Tabela 2.

Campo	Tipo de dados	Descrição
cod_un	String	Código da unidade de gestão
cod_talhao	String	Código do talhão
id_gleba	String	Identificador único do talhão (unidade de gestão x talhão)
ciclo	Integer	Nº do ciclo de plantação
rotacao	Integer	Nº da rotação
dt_referen	Date	Data do início da rotação atual
dt_plant	Date	Data da plantação atual
ocupacao	String	Espécie dominante da plantação
idade_ref	Float	Anos desde o início da rotação
idade_plant	Float	Anos desde o início da plantação
area_ha	Float	Área do talhão (ha)
geometry	MULTIPOLYGON	Tipo de geometria

Tabela 2 - Atributos da camada "NVG_2015-2023_Próprios_clean"

No âmbito deste trabalho, destaca-se o campo "dt_referen" que contém a data em que se iniciou a última rotação de cada talhão, e que por isso permite obter uma estimativa da última vez que o talhão foi cortado.

Tabelas auxiliares

Juntamente com a camada espacial existem outras 2 tabelas sobre as operações de silvicultura e exploração. A tabela de silvicultura não possui informações relevantes para esta parte do trabalho, pelo que não será abordada neste documento.

Tabela de exploração "Exploracao_NVG_2015-2023_Proprios_clean": Esta tabela possui detalhes sobre as várias operações de exploração realizadas, incluindo os eventos de cortes rasos que ocorreram nos povoamentos florestais abrangidos pela base de

dados, as suas datas e os tipos de intervenção. A sua tabela de atributos possui os seguintes campos (Tabela 3):

Campo	Tipo de dados	Descrição
Id Projeto	Integer	Código da unidade de gestão
Talhão	String	Código do talhão
Id Gleba	String	Identificador único (unidade de gestão e talhão)
Data Real	String	Início da operação
Atividade	String	Tipo de operação
Manejo	String	Método utilizado na operação

Tabela 3 - Atributos da tabela de exploração “Exploracao_NVG_2015-2023_Proprios_clean”

O campo “Data Real” fornece a informação sobre as datas das várias operações que foram realizadas no talhão e o campo “Atividade” informa sobre o tipo de operação que foi realizada. A conjugação destas informações permite perceber as datas das operações de cortes rasos que foram realizadas em cada talhão.

Ao contrário da informação expressa no campo “data_referen” da tabela de atributos principal dos dados, que só permite uma entrada por talhão, na tabela de exploração o mesmo talhão pode ter várias entradas, tantas quanto o nº de atividades de exploração realizadas.

Desta forma, o campo “data_referen” é uma data atribuída a todo o talhão, mas que não corresponderá exatamente à data de corte de toda a sua área. Essa informação está expressa na tabela de exploração, mas como esta não contém dimensão espacial, só é possível fazer corresponder uma operação de corte ao talhão correto, não sendo possível obter a informação exata da área que foi cortada.

Metodologia

Problema inicial e abordagem seguida

Dadas as características específicas dos dados, estes exigem um pré-processamento para que possam ser utilizados como validação ou treino de algoritmos de deteção automática de perdas de vegetação.

O problema chave está colocado no facto de não existir informação sobre a área que foi cortada dentro de um talhão em cada uma das operações de corte.

Assim, a metodologia utilizada tem como principal objetivo determinar a área que foi cortada em cada uma das operações registadas na tabela de exploração. Os métodos aplicados podem ser divididos em duas partes:

1. Divisão dos talhões em áreas menores (sub-talhões);
2. Utilização de dados Sentinel-2 para estimar, dentro dos intervalos indicados nos dados originais, a data exata de cada corte de cada sub-talhão.

Divisão dos sub-talhões

A divisão dos talhões em sub-talhões foi baseada nos limites existentes na base de dados original. Esta delimitação segue o desenho original da The Navigator Company e as divisões são baseadas em estradas, caminhos florestais, faixas de gestão de combustíveis ou outros elementos que criem descontinuidades nos povoamentos florestais.

A utilização desta delimitação não permite obter a área exata que foi cortada de uma só vez, mas é a melhor aproximação que é possível de realizar de forma automática.

Atribuição de uma data de corte a cada sub-talhão

A determinação de uma data de corte para cada sub-talhão foi efetuada com base no Índice de Diferença Normalizada de Vegetação (NDVI), calculado a partir de dados Sentinel-2, e permitiu fazer a correspondência entre as datas registadas na tabela de exploração e os sub-talhões. O processo encontra-se descrito nos parágrafos seguintes.

Após o pré-processamento das imagens Sentinel-2, foi efetuado o cálculo do NDVI para cada uma das imagens disponíveis nos sub-talhões, dentro do intervalo definido pelas datas das operações de corte registradas na tabela de exploração. Este intervalo foi determinado com base na última e na primeira data que constavam na tabela, tendo sido subtraídos 60 dias à primeira data e adicionados 60 à última, de forma a alargar a janela temporal.

Em seguida, foi calculado um valor de mediana de NDVI de cada um dos sub-talhões para cada uma das imagens Sentinel-2 disponíveis.

Posteriormente, foi sucessivamente calculada a diferença de mediana entre duas instâncias consecutivas e determinada qual a maior quebra registada na janela temporal considerada.

Para determinar a data a utilizar como referência na BDR foi selecionada a data da operação de corte da tabela de exploração imediatamente anterior à data da maior quebra de NDVI.

O resultado final são as áreas dos dados originais com uma data de corte atribuída a cada sub-talhão com base em dados Sentinel-2, mas respeitando as datas fornecidas pela The Navigator Company.

Tabela de atributos

Os campos da tabela de atributos estão descritos na Tabela 4.

Campos	Tipo	Descrição
Fid	Integer	Identificador único de cada linha da tabela de atributos
Id_gleba	String	Identificador único do talhão
Ocupacao	String	Espécie dominante da vegetação
id	String	Identificador único do sub-talhão (baseado no id_gleba)
First_start_date	String	Data do primeiro corte registado na tabela de exploração

First_end_date	String	Data do último corte registado na tabela de exploração
Nr_empty_cells	String	Nº de datas em que o valor da mediana de NDVI do Sentinel-2 é nula
Nr_non_empty_cells	String	Nº de datas em que o valor da mediana de NDVI do Sentinel-2 não é nula
Nr_S2_dates	String	Nº de datas de Sentinel-2 disponíveis para cada parcela dentro do intervalo de tempo definido
Date_of_biggest_drop	String	Data da maior quebra na mediana de NDVI
First_estimated_date	String	Data da tabela de exploração imediatamente antes da maior quebra de NDVI
Area_ha_su	Float	Área do sub-talhão
geometry	POLYGON	Coordenadas do singlepart no EPSG:3763 ETRS89-PT-TM06

Tabela 4 - Atributos da BDR final

O campo “First_estimated_date” expressa a data que deve ser considerada aquando da utilização dos dados para validar produtos de perdas de vegetação.

Limitações da BDR

A forma como a BDR foi construída apresenta algumas limitações que se encontram listadas abaixo:

1. A divisão dos sub-talhões não foi baseada nas áreas que foram cortadas em cada operação de exploração, mas sim nas fronteiras definidas pelo desenho dos polígonos. Desta forma, a área de um sub-talhão pode não ter sido cortada num único evento, pelo que a data que lhe foi atribuída pode não ser válida para toda a sua área.
2. Em alguns casos, o nº ou frequência de imagens Sentinel-2 disponíveis não foi considerado suficiente para a determinação da maior quebra de NDVI no

sub-talhão. Nessas situações, os campos “Date_of_biggest_drop” e “First_estimated_date” ficaram sem valor.

3. Em alguns sub-talhões a data da maior quebra de NDVI é anterior à primeira data de corte registada na tabela de exploração. Nestes casos, a opção foi a de deixar vazio o campo “First_estimated_date”, não existindo uma data exata para utilizar na validação.