E2.1 – Report with the methodology for creating the Reference Database

Table of Contents

1 Reference Database

1.1 Introduction

The main goal of the Navigator reference database within the current project is to be used as reference and validation to train detection and classification methods to predict vegetation losses over Continental Portugal from Sentinel-2 image collections.

In this report, we first describe the data base contents and range. Then, we describe how the database was structured to avoid redundancy and assure consistency.

Given the project's goals, it is necessary that the reference data base provides a reliable spatial and temporal match between actual changes (like clear-cuts) and reported change dates within the data base. However, that was not the case, since there was a significant uncertainty on the actual date of field interventions (like clear-cuts) when the parcel is large and the contains multiple sub-parcels. To solve that problem, we relied on Sentinel-2 image collections to estimate the approximate date of vegetation loss for each sub-parcel, and we matched that estimated date to the closest date available in the database for the corresponding parcel.

From the structured data base, and incorporating the estimates of intervention dates at the sub-parcel level, we created a single chronologically unified table, with traces for every sub-parcel in the database, all the reported field interventions and their respective dates. This final table has the correct format to be readily used as reference for training and validation, according to the project's goals.

Python scripts related to this project task are available at https://github.com/manuelcampagnolo/S2CHANGE/tree/main/scripts/BDR

1.2 Database Content

The dataset to be utilized in this project was provided to us by the Navigator Company. The dataset primarily consists of geospatial information encapsulated within a geopackage format. This geopackage includes several key components, including a polygon shapefile delineating forested areas in Continental Portugal, and two tabular datasets to complement the geospatial information.

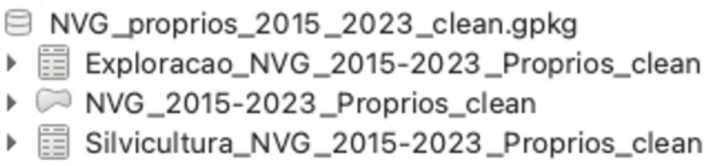

NVG_proprios_2015_2023_clean.gpkg
Exploracao_NVG_2015-2023_Proprios_clean
NVG_2015-2023_Proprios_clean
Silvicultura_NVG_2015-2023_Proprios_clean

Figure 1 - NVG_proprios_2015_2023_clean geopackage

The dataset includes over 27,000 hectares of forested land across Portugal, with Beira Baixa (8542 ha), Alentejo Litoral (7040 ha), and Algarve (2318 ha) collectively accounting for over 60% of this expanse

1.2.1 NVG_2015_2023_Proprios_clean

The shapefile contains polygons representing forest areas. Each polygon has information regarding plantation dates, cycle, rotation, and types of occupation, as shown in Table 1.

Table 1 - Metadata of NVG_2015_2023_Proprios_clean table

Attribute	**Description**
cod_un	Management unit code
cod_talhao	Parcel code
id_gleba	Management unit code + parcel code
ciclo	Plantation cycle
rotacao	Nr of rotation within the plantation cycle
dt_referen	Start date of rotation
dt_plant	Start date of the plantation cycle
ocupacao	Dominant species and form of plantation
idade_ref	Years since reference date (age of trees)
idade_plant	Years since plantation date (plantation cycle age)
area_ha	Parcel's area in hectares
geometry	feature's geometry of type MULTIPOLYGON

1.2.2 Exploração_NVG_2015-2023_Proprios_clean

Tabular data with details about clear-cut events that have occurred within the forest parcels. Each entry includes the operations dates and description. The operation in this table includes clear-cuts (corte c/ casca; corte s/ casca) and transport of the cut wood (rechega c/casca; rechega s/ casca). The metadata is shown on Table 2.

Table 2 - Metadata of Exploração_NVG_2015_2023_Proprios_clean table

Attribute	**Description**
Id Projeto	Management unit code
Talhão	Parcel code
Id Gleba	Management unit code + parcel code

Data Real	Operation start date
Atividade	Type of operation
Manejo	Operation management

1.2.3 Silvicultura_NVG_2015-2023_Proprios_clean

Tabular data that provides comprehensive information on forestry operations events conducted within forest parcels. Encompassing a variety of operations, the dataset delineates activities like planting (plantação), harrowing (gradagem) plowing (ripagem), retancha, rod selection (selecção de varas), and clearing undergrowth (limpeza de mato – manual ou mecânica). The metadata is shown on Table 3.

Table 3 - Metadata of Silvicultura_NVG_2015_2023_Proprios_clean table

Attribute	**Description**
Id	
Data Operação	Forestry operation start date
Id Projeto	Management unit code
Talhão	Parcel code
Id_gleba	Management unit code + parcel code
Desc. Atividade	Type of forestry operation

Figure 2 shows an example of clear-cut events and forestry operations done in one plot. These operations were pinpointed against Sentinel-2 NDVI temporal series from august 2018 until December 2021.

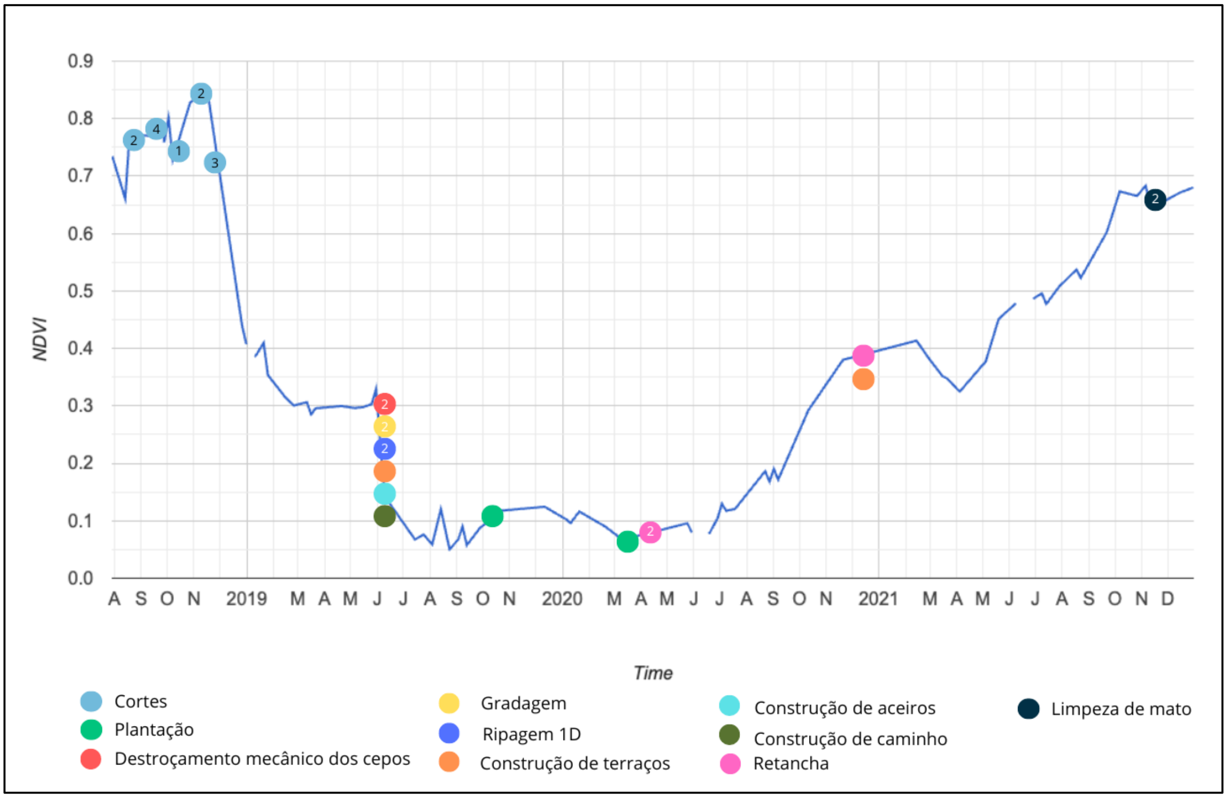


Figure 2 – Time series of NDVI for plot 56004-T010_EG. Colored circles represent clear-cut events and forestry activities on the plot, with numbers indicating the frequency of these activities. For instance, the initial blue circle labeled "2" means two clear-cut dates in August.

The dataset is organized into hierarchical management units, each further subdivided into parcels, as shown in Figure 3.

1.2.4 Management Units

Management units (Unidades de Gestão) serve as the primary divisions within the dataset. Each management unit is assigned a unique identifier in the form of a code, enabling straightforward identification, and referencing. The dataset has 531 management units.

1.2.5 Parcels

Parcels (Talhões) are finer subdivisions within each management unit in the dataset. Each parcel is uniquely identified by a combination of the management unit code and the parcel code, referred to as 'id_gleba'. Parcels are grouped based on the same land management type.

In total, there are 1583 parcels represented by polygons in the shapefile. Notably, 1383 of these polygons are multipart, indicating that each feature is composed of more than one physical part, yet it only references one set of attributes. Figure 3 shows an example of the organization of the dataset.

1.2.6 Sub-parcels

Although sub-parcels (sub-talhões) are not currently integrated into the initial database, they serve as a layer for potential data expansion. They can be seen as a mean to further refine the spatial organization within parcels, facilitating more detailed analysis. The unique identifier for these sub-parcels, labeled as "id", was generated by appending an incremental value to the existing parcel identifier ("id_gleba").

Figure 3 illustrates the hierarchical structure of the database organization. Taking management unit 50445 as an example, three levels are observed. The first level represents management units – identified by "cod_un" –, the second level comprises parcels – identified by "id_gleba" -, and the third level contains the sub-parcels – identified by "id". For instance, parcel T001_EG contains 19 sub-parcels, denoted by the unique identifier 50445-T001_EG_{i}, ranging from 01 to 19, while parcel T002_EG has 14 sub-parcels, with identifiers ranging from 01 to 14.

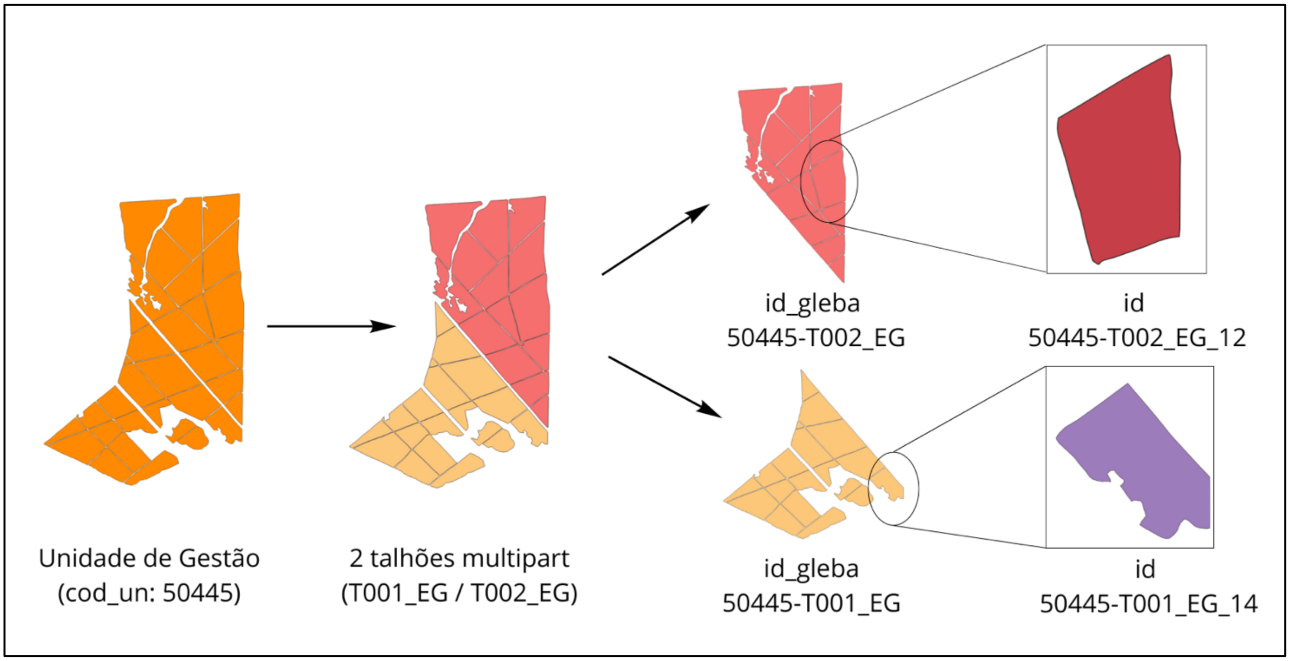


Figure 3 - Hierarchical structure of the database using two parcels (50445-T001_EG and 50445-T002_EG) as examples. Represented in orange is a single management unit (unidade de gestão) subdivided into two parcels (talhões) highlighted in yellow and pink. These parcels are further delineated into sub-parcels (sub-talhões) shown in red and purple.

1.3 Geographic Distribution

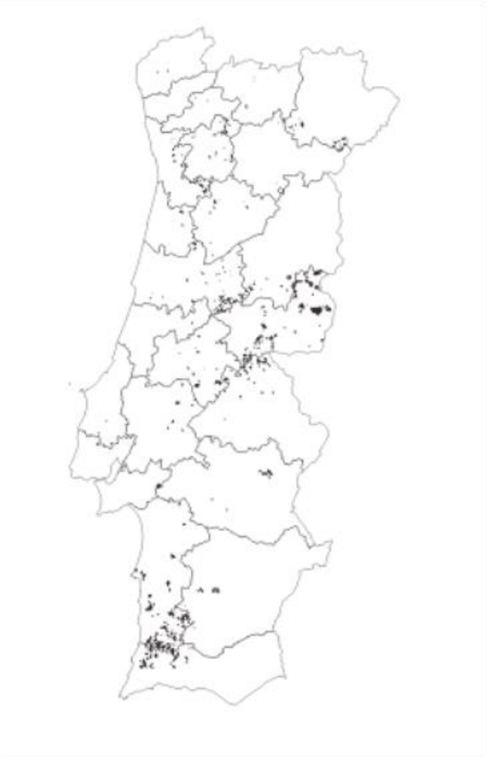


Figure 4 - Geographic distribution of the database in mainland Portugal

The dataset comprises over 27,000 hectares of forested land across Mainland Portugal. Geographically, more than 60% of this area is distributed between Beira Baixa (8542 ha), Alentejo Litoral (7040 ha), and Algarve (2318 ha) combined. Conversely, the less prominent NUTs III regions together account for less than 5% of the total forested area.

1.3.1. Dominant Forest Species

The dataset has 531 management units and 1583 parcels covering 27439 hectares of forest management. There are 4 classes of species: Eucalipto, Pinheiro-manso, Pinheiro-bravo, and 'Outros Pinheiros'. Table 4 shows the area distribution where Eucaliptos Forest represents 92,5% of the total forest area with 25372 hectares, while Pinheiro-manso, Pinheiro-bravo, and Outros Pinheiros represent 1,77%, 3,87%, and 1,9%, with 484, 1062 and 523 hectares, respectively.

In terms of parcel distribution, Eucalipto Forest comprises the majority with 1408 parcels, followed by Pinheiro-bravo with 121 parcels, Pinheiro-manso with 35 parcels, and Outros Pinheiros trees with only 19 parcels.

Table 5 outlines the average parcel sizes for each vegetation class, with 'Outros Pinheiros' parcels averaging 27 hectares, Eucalipto parcels averaging 18 hectares, followed by Pinheiro-manso with 13 hectares and Pinheiro-bravo with 9 hectares.

Table 4 - Area in hectares, number of parcels and respective percentages per vegetation class

Vegetation Class	Area (ha)	Percentage (%)	Nr of parcels	Percentage (%)
Eucalipto	25370,5	92,46	1408	88,95
Pinheiro-manso	484,4	1,77	35	2,21
Pinheiro-bravo	1061,9	3,87	121	7,64
Outros Pinheiros	522,5	1,9	19	1,2
Total	**27439,31**	**100**	**1583**	**100**

Table 5 - Average, minimum and maximum areas of parcels per vegetation class

Vegetation Class	Average of parcel area (ha)	Minimum of parcel area (ha)	Maximum of parcel area (ha)
Eucalipto	18,02	0,022	286
Pinheiro-manso	13,84	0,042	252,58
Pinheiro-bravo	8,78	0,008	289,65
Outros Pinheiros	27,5	0,089	114,28

As mentioned before, most of the parcels are multipart polygons. In 1583 features, only 200 are singlepart (represented by one polygon only). The Eucalipto class, being the most well-represented area-wise, has more total parcels – singlepart and multipart. Table 6 represents the total number of parcels and their distribution between single and multipart, while Table 7 considers multipart parcels and shows total number, average, minimum and maximum of sub-parcels.

Table 6 - Total number of parcels, number of single and multipart parcels per vegetation class

Vegetation Class	Total number of parcels	Total number of single part parcels	Total number of multipart parcels
Eucalipto	1408	140	1268
Pinheiro-manso	35	9	26
Pinheiro-bravo	121	46	75
Outros Pinheiros	19	5	14
Total	**1583**	**200**	**1383**

Table 7 - Considering multipart parcels. Total, average, minimum, and maximum number of sub-parcels per vegetation class.

Vegetation Class	**Total number of sub-parcels**	**Average number of sub-parcels**	**Minimum number of sub-parcels**	**Maximum number of sub-parcels**
Eucalipto	18653	13	2	377
Pinheiro-manso	249	7	2	51
Pinheiro-bravo	641	5	2	69
Outros Pinheiros	159	8	2	34
	19702			

Considering a minimum parcel area of 0.5 hectares, Table 8 provides the total parcel count and their respective areas, in hectares, categorized by vegetation class.

Table 8 - Total number of parcels, per vegetation class, with an area in hectares higher than 0,5.

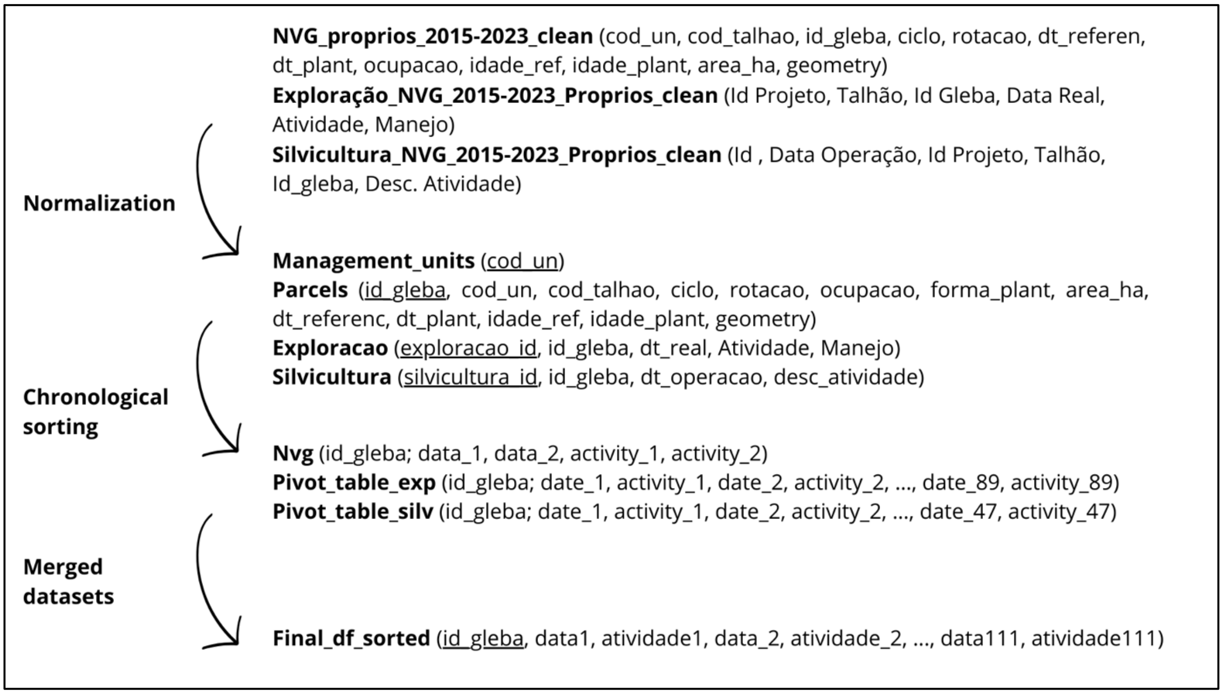
Vegetation Class	**Total number of parcels**	**Total area (ha)**
Eucalipto	1341	25352.56
Pinheiro-manso	27	482.70
Pinheiro-bravo	85	1053.31
Outros Pinheiros	15	521.61
	1468	**27410.19**

2 Methodology

2.1 Database organization

The collected data underwent preprocessing steps to standardize and harmonize its format. These procedures were implemented utilizing Python within QGIS, primarily leveraging the Pandas library.

Figure 5 shows the main step of preprocessing until the final dataset, sorted chronologically.



NVG_proprios_2015-2023_clean (cod_un, cod_talhao, id_gleba, ciclo, rotacao, dt_referen, dt_plant, ocupacao, idade_ref, idade_plant, area_ha, geometry)
Exploração_NVG_2015-2023_Proprios_clean (Id Projeto, Talhão, Id Gleba, Data Real, Atividade, Manejo)
Silvicultura_NVG_2015-2023_Proprios_clean (Id , Data Operação, Id Projeto, Talhão, Id_gleba, Desc. Atividade)

Normalization

Management_units (cod_un)
Parcels (id_gleba, cod_un, cod_talhao, ciclo, rotacao, ocupacao, forma_plant, area_ha, dt_referenc, dt_plant, idade_ref, idade_plant, geometry)
Exploracao (exploracao_id, id_gleba, dt_real, Atividade, Manejo)
Silvicultura (silvicultura_id, id_gleba, dt_operacao, desc_atividade)

Chronological sorting

Nvg (id_gleba; data_1, data_2, activity_1, activity_2)
Pivot_table_exp (id_gleba; date_1, activity_1, date_2, activity_2, ..., date_89, activity_89)
Pivot_table_silv (id_gleba; date_1, activity_1, date_2, activity_2, ..., date_47, activity_47)

Merged datasets

Final_df_sorted (id_gleba, data1, atividade1, data_2, atividade_2, ..., data111, atividade111)

Figure 5 - Main steps of data preprocessing until final dataframe. Starting with normalization of the initial tables, sorting them chronologically and merging the datasets to obtain the final_df_sorted.

2.1.1 Renaming columns

To ensure consistency, columns and table names were adjusted. These adjustments involved the removal of white spaces and special characters (Table 9).

Table 9 - Column names before and after renaming columns

Original name	***New name***
Exploração_NVG_2015-2023_Proprios_clean	
Exploração_NVG_2015-2023_Proprios_clean	Exploracao_NVG_2015-2023_Proprios_clean
Id Projeto	cod_un
Talhão	cod_talhao
Id Gleba	id_gleba
Data Real	dt_real
Silvicultura_NVG_2015-2023_Proprios_clean	
Id Projeto	cod_un
Talhão	cod_talhao
Data Operação	dt_operacao
Desc. Atividade	desc_atividade

2.1.2 Normalization

2.1.2.1What is data normalization?

Database normalization is the process of structuring a database according to what's called **normal forms**, with the final product being a relational database, free from data redundancy. The main objective of database normalization is to eliminate redundant data, minimize data modification errors, and simplify the query process. This process is essential to deal with large amounts of data. Normalization organizes columns and tables of a database according to a set of **normal form rules** – the process aims to ensure that if any data is updated, inserted, or deleted, the integrity of the database stays intact (Date, 2019).

2.1.2.2Primary key and Foreign key

Each normalized table should have a **primary key**. A primary key serve as a unique identifier for each record in a table, ensuring that each row is distinct and identifiable. By having a primary key, data integrity is maintained, as it prevents duplicate records and ensures that each row can be uniquely identified (Hernandez, 2021)

On the other hand, **foreign keys** establish relationships between tables by linking a column in one table to the primary key of another table. This linkage maintains referential integrity, ensuring that relationships between tables are preserved and data consistency is maintained. This key is a minimum set of attributes in a relation that matches the primary key of another relation.

2.1.2.3Normal Forms

A table is considered normalized when it reaches the third normal form (Date, 2019)

- First Normal Form (1NF) – a relation is in 1NF if each table cell should contain only a single value, and each column should have a unique name.
- Second Normal Form (2NF) - builds upon 1NF and requires that each non-key attribute is dependent on the primary key. This means that each column should be fully functionally dependent on the primary key
- Third Normal Form (3NF) - builds upon 2NF by requiring that all non-key attributes be independent. This means each column should be directly related to the primary key, not to any other columns in the same table.

2.1.2.3.1 NVG Table

The 'ocupacao' column, which originally held information about the dominant species and the plantation method, has now been divided into two separate columns. One column, 'ocupacao', contains data regarding the dominant tree species, while the other column, 'forma_plant', specifies the plantation method.

This table originated two tables. The primary key on table “Management units” is ‘cod_un’, and for “Parcels” table is ‘id_gleba’.

Management_units (cod_un)

Parcels (id_gleba, cod_un, cod_talhao, ciclo, rotacao, ocupacao, forma_plant, area_ha, dt_referenc, dt_plant, idade_ref, idade_plant, geometry)

2.1.2.3.2 Exploracao and Silvicultura Tables

For tables Exploracao and Silvicultura, a new attribute was created to act as a primary key, leaving id_gleba as a foreign key. This attribute is the name of the table with the suffix ‘_id’.

Exploracao (exploracao_id, id_gleba, dt_real, Atividade, Manejo)

Silvicultura (silvicultura_id, id_gleba, dt_operacao, desc_atividade)

Note that the relation between table Parcels and table Exploracao is of type 1:n, which means that a single feature of Parcels can be associated to several events in Exploracao. The same occurs between tables Parcels and Silvicultura.

The number of duplicates is:

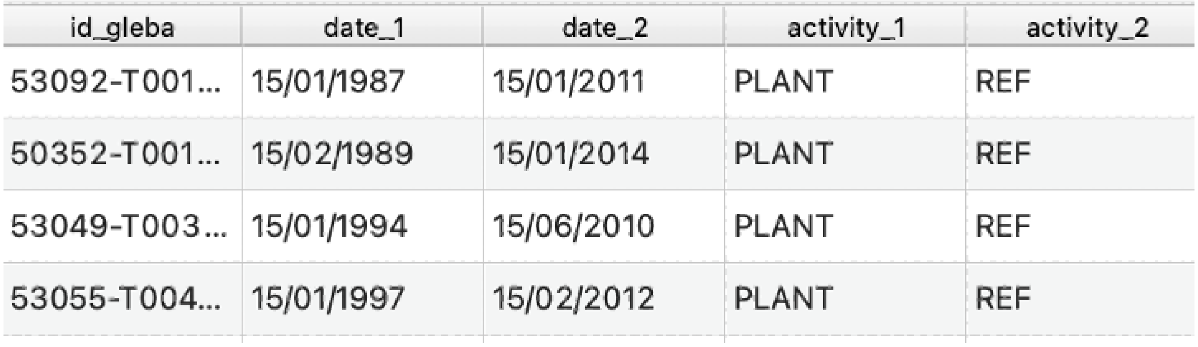
- Number of duplicate rows from exploracao table: 483
- Number of duplicate rows from silvicultura table: 1001

2.2 Chronologically unified table

To create the database with chronological sorting, it's essential for all tables to maintain uniformity in format. Here, only the columns of dates and descriptions are of interest.

The data from the Parcels table was parsed to extract the 'dt_referenc' and 'dt_plant' columns into a new data frame. Additionally, a descriptive column named 'REF' was appended to denote reference data and 'PLANT' for planting date. As the original, the resultant data frame contains 1583 entries, yet comprising only 4 columns: 'data_1', 'data_2', 'activity_1', and 'activity_2'. Table 10 shows, as an example, the first 4 rows of table NVG.

Table 10 - First four rows of NVG table. This table has 5 columns in total considering only the planting and reference dates, as well as their description – "PLANT" for planting and "REF" for reference date



id_gleba	date_1	date_2	activity_1	activity_2
53092-T001...	15/01/1987	15/01/2011	PLANT	REF
50352-T001...	15/02/1989	15/01/2014	PLANT	REF
53049-T003...	15/01/1994	15/06/2010	PLANT	REF
53055-T004...	15/01/1997	15/02/2012	PLANT	REF

The procedures followed for both the 'exploracao' and 'silvicultura' tables mirrored each other. These tables underwent a transformation into a pivot table format, utilizing columns for dates and descriptions, with 'id_gleba' acting as the primary key. Furthermore, chronological sorting was applied to both tables.

Three intermediate tables were created – Nvg, Pivot_table_exp, Pivot_table_silv. Subsequently, the three tables were merged to generate the final dataset – **Final_df_sorted**. Table 11 shows, as an example, the first 4 rows of final_df_sorted table.

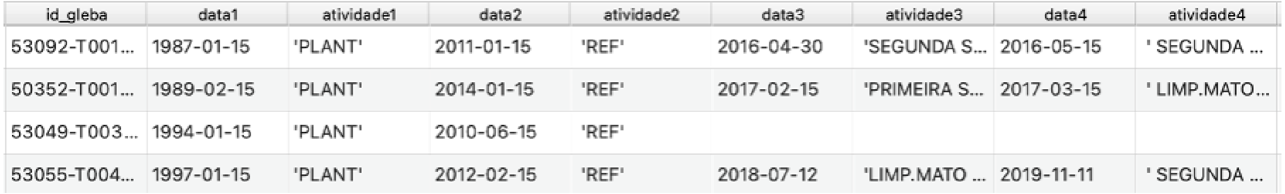
Nvg (id_gleba; data_1, data_2, activity_1, activity_2)

Pivot_table_exp (id_gleba; date_1; activity_1; …; date_89; activity_89)

Pivot_table_silv (id_gleba; date_1; activity_1; …; date_47; activity_47)

Final_df_sorted (id_gleba, data1, atividade1, ..., data111, atividade111)

Table 11 – First four rows of the final_df_sorted. Dates are organized by chronological order with their corresponding activities



id_gleba	data1	atividade1	data2	atividade2	data3	atividade3	data4	atividade4
53092-T001...	1987-01-15	'PLANT'	2011-01-15	'REF'	2016-04-30	'SEGUNDA S...	2016-05-15	' SEGUNDA ...
50352-T001...	1989-02-15	'PLANT'	2014-01-15	'REF'	2017-02-15	'PRIMEIRA S...	2017-03-15	' LIMP.MATO...
53049-T003...	1994-01-15	'PLANT'	2010-06-15	'REF'				
53055-T004...	1997-01-15	'PLANT'	2012-02-15	'REF'	2018-07-12	'LIMP.MATO ...	2019-11-11	' SEGUNDA ...

2.3 Estimated clear-cut dates for sub-parcels

The main goal of the Navigator reference database within the current project is to be used as reference and validation to train detection and classification methods to predict vegetation losses over Continental Portugal from Sentinel-2 image collections. Towards that end, it is necessary that the reference data base provides a reliable match between actual changes and reported change dates within the data base.

However, the main structural unit of the database is the parcel even though activities like clear cuts are performed at sub-parcel level. Therefore, dates in the database, which we assume that are actual clear-cut dates, are not assigned to a specific sub-parcel. This is illustrated in Figure 6 below, that also depicts Sentinel-2 images of Normalized Difference Vegetation Index (NDVI), and shows clear-cuts distributed along a range of 45 days for a given parcel. The original database contains a series of dates for clear-cuts for the parcel but does not associate each of those dates to a particular sub-parcel. Therefore, we face what is known as a “matching problem” between the set of sub-parcels and the set of candidate clear-cut dates that are available in the original reference database.

To solve this relevant shortcoming of the database provided by the Navigator company, we relied on series of Sentinel-2 NDVI. The idea was to create for each sub-parcel the time series of the median NDVI value and determine the date of the maximum drop in NDVI. Then, we assigned the closest clear-cut date in the database for that same parcel to the sub-parcel. The procedure is illustrated below with a concrete example.

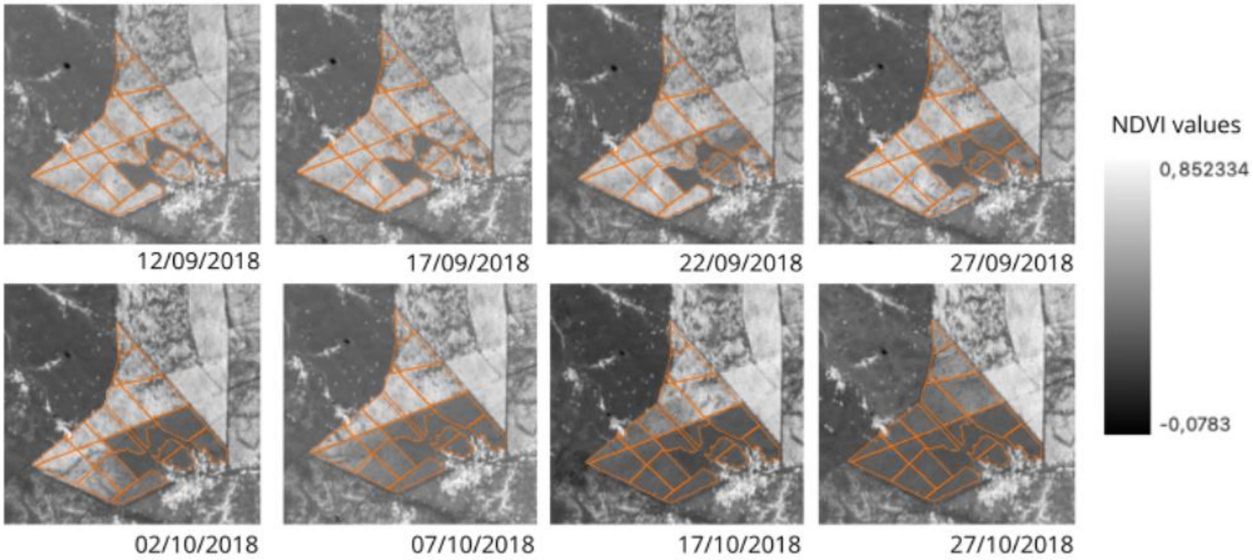


Figure 6 - NDVI images from september to october of 2018. The represented polygons are the parcel with 'id_gleba' = 50445-T001_EG. Low NDVI values indicate a low vegetation content meaning that a clear-cut might have happened. These are represented by a darker color.

To assign clear-cut dates to individual sub-parcels within parcels, a conversion to singlepart polygons is necessary. This transformation enables the treatment of each sub-parcel as a distinct unit.

To achieve this, the next steps are followed:

1. Conversion of the NVG shapefile to singlepart, using “multipart to singlepart” tool on QGIS.

2. Retrieval of the NDVI median values within the temporal series of interest, using Google Earth Engine (GEE) API, and storing them in a csv file
3. From the csv file, create a pivot table organized by the 'id' of the sub-parcel.
4. Calculation of the most significant drop in NDVI value and identification of the corresponding date.
5. Comparison of the resulting clear-cut dates from GEE with those stored in the chronological database.

2.3.1 NDVI Median

Using the parcel example (id_gleba = 50445-T001_EG), Figure 7 illustrates the median NDVI trend over the months. To establish the start and end dates for extracting NDVI values, two months before and after the first and last clear-cut dates were considered. To identify more than one clear-cut event per parcel, the condition of an interval of 2 years between consecutive dates was considered.

First clear-cut date = 2018-09-17
- 2 months
Start date = 2018-07-17

Last clear-cut date = 2018-10-29
+ 2 months
End date = 2018-12-29

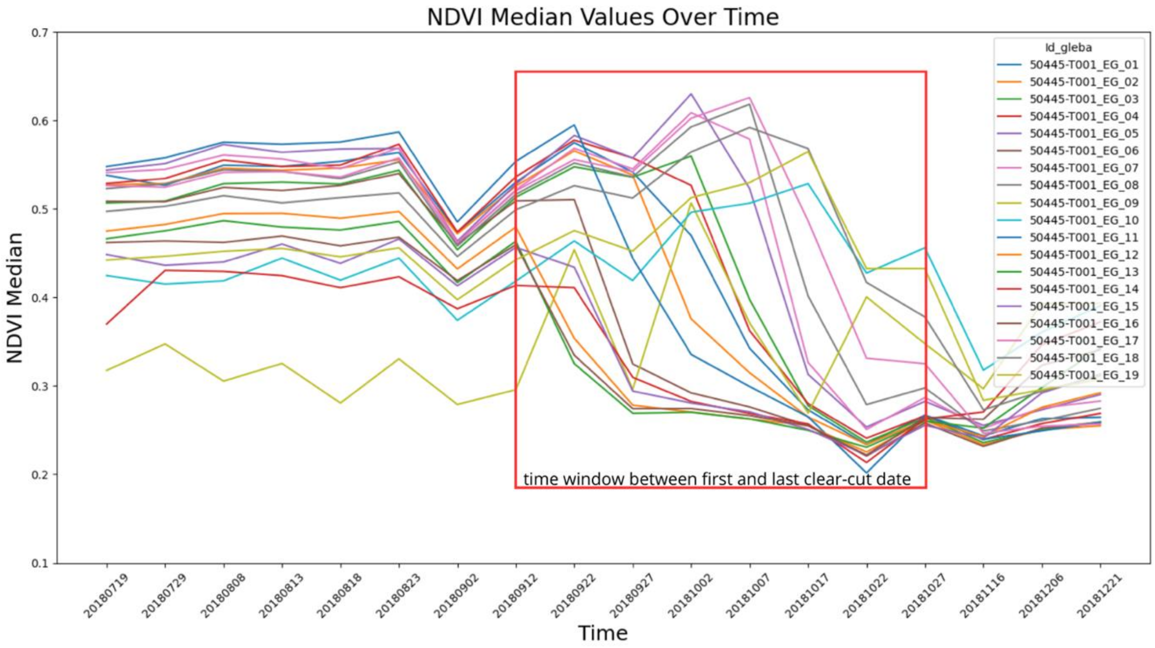


Figure 7 - Median of NDVI over time of each sub-parcel of parcel with id_gleba = 50445-T001_EG. Each line represents one sub-parcel, and the red square indicates the time window between the first and last clear-cut date in our sorted chronologically database.

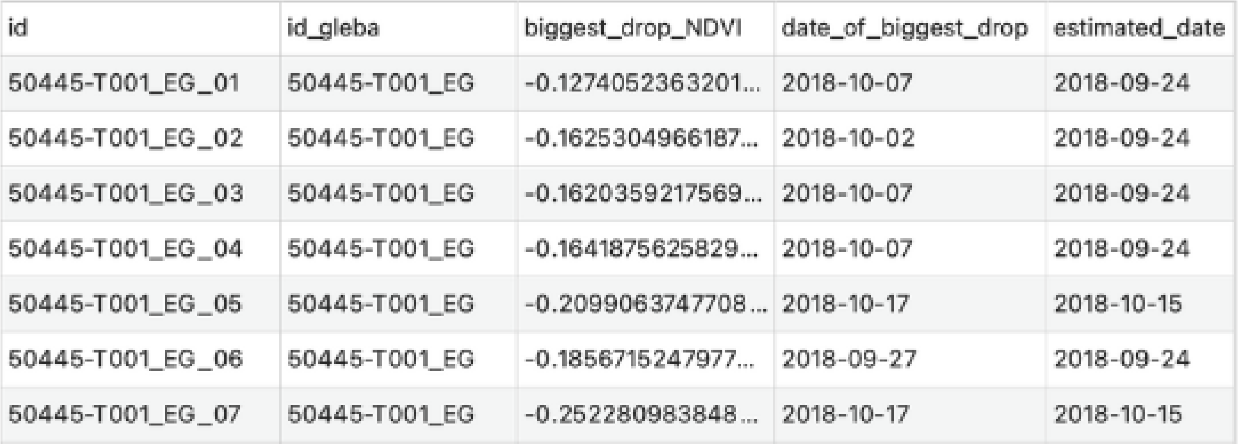
2.3.2 Largest NDVI drop and estimated clear-cut dates

The largest decrease in NDVI value was determined by comparing the median NDVI values for each sub-parcel across consecutive dates and identifying the two dates with the most significant change.

Based on the dates provided for each parcel alongside the dates of the most significant NDVI decline for each sub-parcel, we were able to project the clear-cut dates for the sub-parcels. The dates of the maximum drop in NDVI were sourced from GEE. By comparing these dates with those in the final_df_sorted, it is possible to allocate an estimation of clear-cut date to each sub-parcel considering the closest date before of the biggest drop of NDVI values. Table 12 shows the first rows of the sub-parcels estimated clear-cut dates.

As an example, the first row refers to sub-parcel 01. The biggest NDVI drop happens on October 7th, 2018, indicating a high likelihood of a clear-cut event before that day. The closest date to that day in our final dataset is September 24th, 2018.

Table 12 - First seven rows of pivot table where the largest drop of NDVI was calculated and the clear-cut date were estimated, per sub-parcel. The biggest drop was determined by comparing the median NDVI values across consecutive dates, retriving a date for the biggest drop and compare them with the dates in the final_df_sorted.



id	id_gleba	biggest_drop_NDVI	date_of_biggest_drop	estimated_date
50445-T001_EG_01	50445-T001_EG	-0.1274052363201...	2018-10-07	2018-09-24
50445-T001_EG_02	50445-T001_EG	-0.1625304966187...	2018-10-02	2018-09-24
50445-T001_EG_03	50445-T001_EG	-0.1620359217569...	2018-10-07	2018-09-24
50445-T001_EG_04	50445-T001_EG	-0.1641875625829...	2018-10-07	2018-09-24
50445-T001_EG_05	50445-T001_EG	-0.2099063747708...	2018-10-17	2018-10-15
50445-T001_EG_06	50445-T001_EG	-0.1856715247977...	2018-09-27	2018-09-24
50445-T001_EG_07	50445-T001_EG	-0.252280983848...	2018-10-17	2018-10-15

Finally, the sub-parcels are represented on a map, with their estimated clear-cut dates indicated by colors (Figure 8). Each polygon, corresponding to a specific sub-parcel, is labeled with the date of the estimated clear-cut event, allowing for an easy interpretation of the temporal and spatial distribution of these activities. Moreover, an additional column is added to the chronologically sorted table for each date, labeled "data_estimada{i}", which is updated with a value of "1" if the date matches the estimated clear-cut date shown in the pivot table (see Table 13).

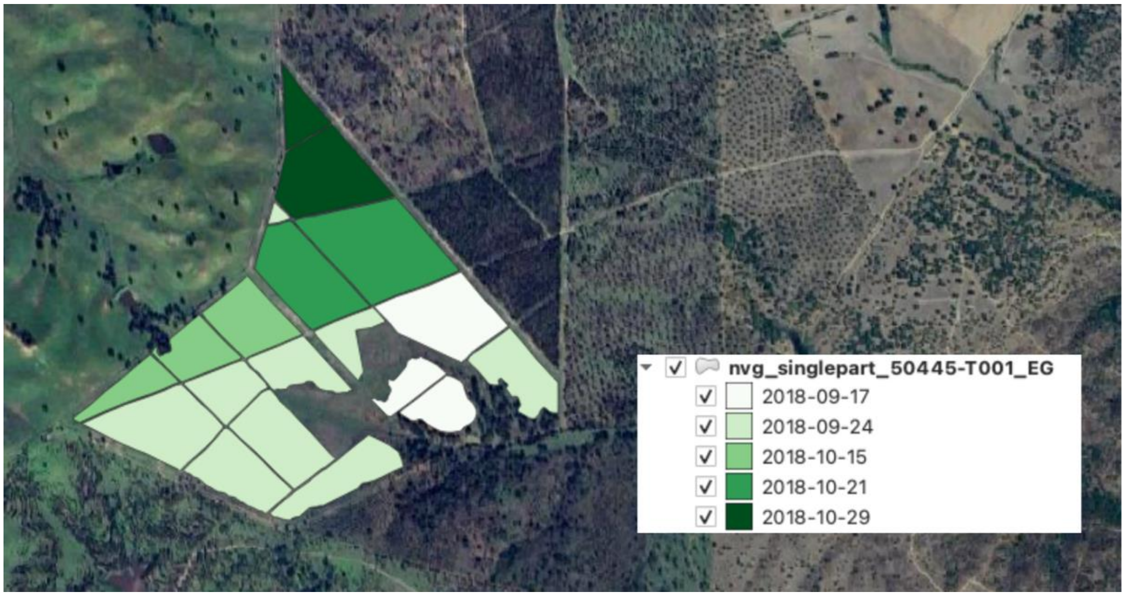
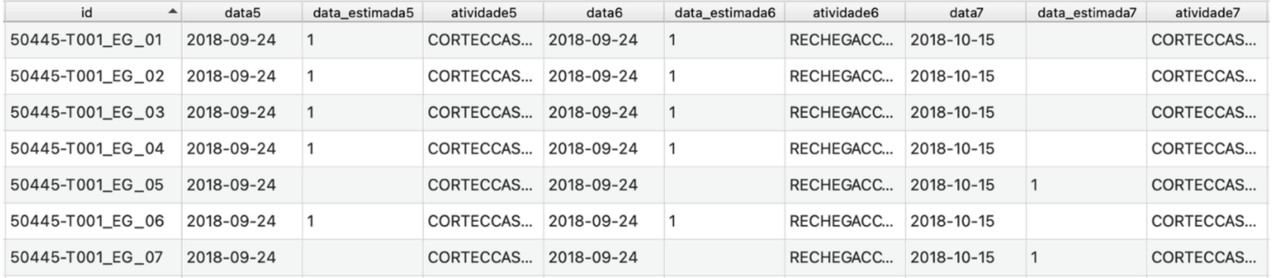


Figure 8 - Temporal distribution of the estimated clear-cut event within the parcel. A clear-cut date was associated to each sub-parcel and labelled with an according color.

Table 13 – First seven rows of the final_df_sorted with and additional “data_estimada” column per “data” column. The extra columns were filled with “1” when the “estimated_date” columns from the pivot table (table12) matched the data column on the final_df_sorted.



id	data5	data_estimada5	atividade5	data6	data_estimada6	atividade6	data7	data_estimada7	atividade7
50445-T001_EG_01	2018-09-24	1	CORTECCAS...	2018-09-24	1	RECHEGACC...	2018-10-15		CORTECCAS...
50445-T001_EG_02	2018-09-24	1	CORTECCAS...	2018-09-24	1	RECHEGACC...	2018-10-15		CORTECCAS...
50445-T001_EG_03	2018-09-24	1	CORTECCAS...	2018-09-24	1	RECHEGACC...	2018-10-15		CORTECCAS...
50445-T001_EG_04	2018-09-24	1	CORTECCAS...	2018-09-24	1	RECHEGACC...	2018-10-15		CORTECCAS...
50445-T001_EG_05	2018-09-24		CORTECCAS...	2018-09-24		RECHEGACC...	2018-10-15	1	CORTECCAS...
50445-T001_EG_06	2018-09-24	1	CORTECCAS...	2018-09-24	1	RECHEGACC...	2018-10-15		CORTECCAS...
50445-T001_EG_07	2018-09-24		CORTECCAS...	2018-09-24		RECHEGACC...	2018-10-15	1	CORTECCAS...

3 References

Date, C. J. (2019). *Database Design and Relational Theory: Normal Forms and All That Jazz*. Apress. https://doi.org/10.1007/978-1-4842-5540-7

Hernandez, M. J. (2021). *Database Design for Mere Mortals®: 25th Anniversary Edition*.