ISA

Development of a forest clear-cut reference database for Continental Portugal at Sentinel-2 resolution

MASTER'S THESIS

Supervisor: Prof. Manuel Campagnolo

U LISBOA | UNIVERSIDADE DE LISBOA

INSTITUTO SUPERIOR DE AGRONOMIA
Universidade de Lisboa
Hinc Patriam Sustinet

Inês Silveira

S2CHANGE - PROJECT

Projeto Agenda Transform (Agenda para a transformação digital das cadeias de valor florestais numa economia portuguesa mais resiliente e hipocarbónica) no subprojecto «P1.5- Dados de deteção remota para a gestão florestal»

What ?

Production of vegetation loss maps based on Sentinel-2 temporal series.

How ?

Change detection techniques - temporal component (pixel level)

Classification techniques - spatial component (patch level)

Why ?

To improve and automate methodologies for the detection and classification of land cover changes in forest and shrub areas, every 2 months.

THESIS OBJECTIVE

What ?

Create a reference database to serve as training data for a change detection algorithm that automates the creation of vegetation loss maps, particularly clear-cut maps, using Sentinel-2 time series data.

How ?

Using a semi-automatic method

Why ?

Accurate and timely generation of vegetation loss maps, contributing to improved monitoring and analysis of land cover changes.

DATA SCIENCE THESIS

WHY?

1. Uses data to answer a question

2. Data-driven workflow
   - Collect and understand the data
   - Prepair and model data
   - Validate and interpret results

3. Use of specialized tools and techniques
   - Programming language: Python and JavaScript
   - Libraries: pandas, geopandas, processing, ...
   - Visualization tools: matplotlib

DATA DESCRIPTION

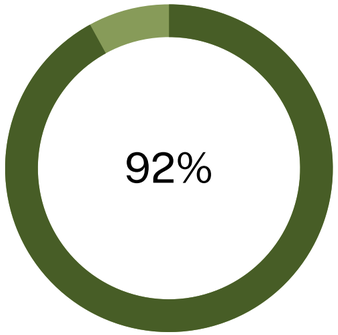
ORIGINAL DATABASE

What do we have?

Mainland Portugal

27.439

Hectares



Eucalyptus

Plantation

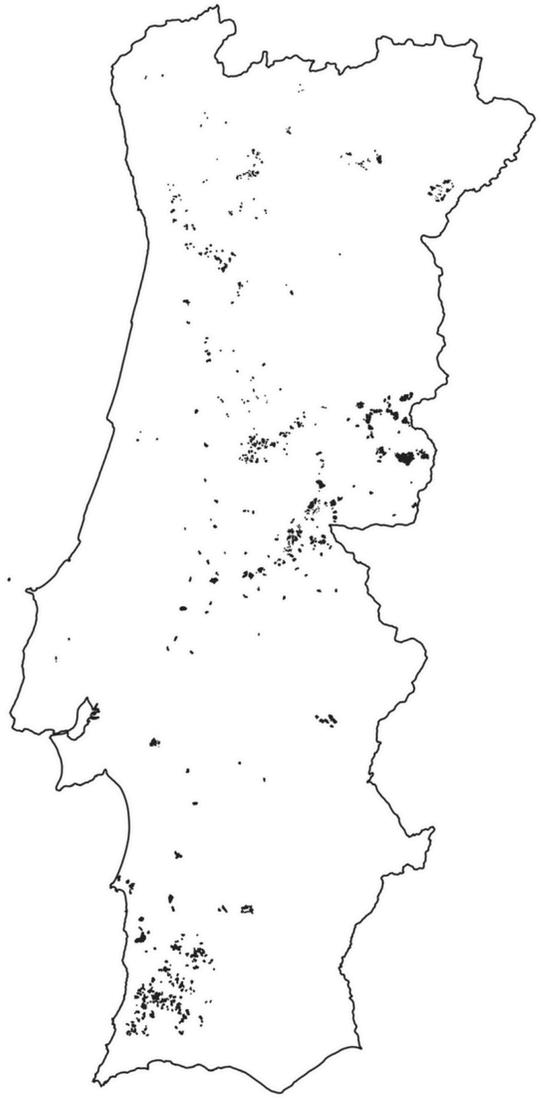
Clear-cuts

Forestry operation

DATA DESCRIPTION

ORIGINAL DATABASE

Geographic distribution



Total

1583 parcels
19702 sub-parcels

Eucalyptus

1408 parcels
18653 sub-parcels

Pinheiro-manso

35 parcels
249 sub-parcels

Pinheiro-bravo

121 parcels
641 sub-parcels

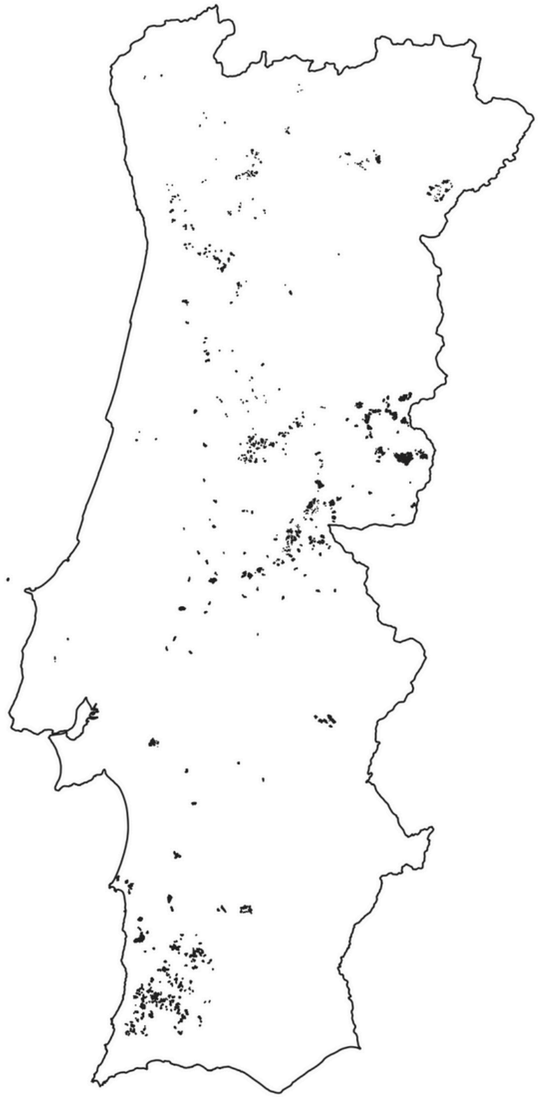
Other pine trees

19 parcels
159 sub-parcels

DATA DESCRIPTION

ORIGINAL DATABASE

Geographic distribution



Total

1583 parcels
19702 sub-parcels

Eucalyptus

1408 parcels
18653 sub-parcels

Pinheiro-manso

35 parcels
249 sub-parcels

Pinheiro-bravo

121 parcels
641 sub-parcels

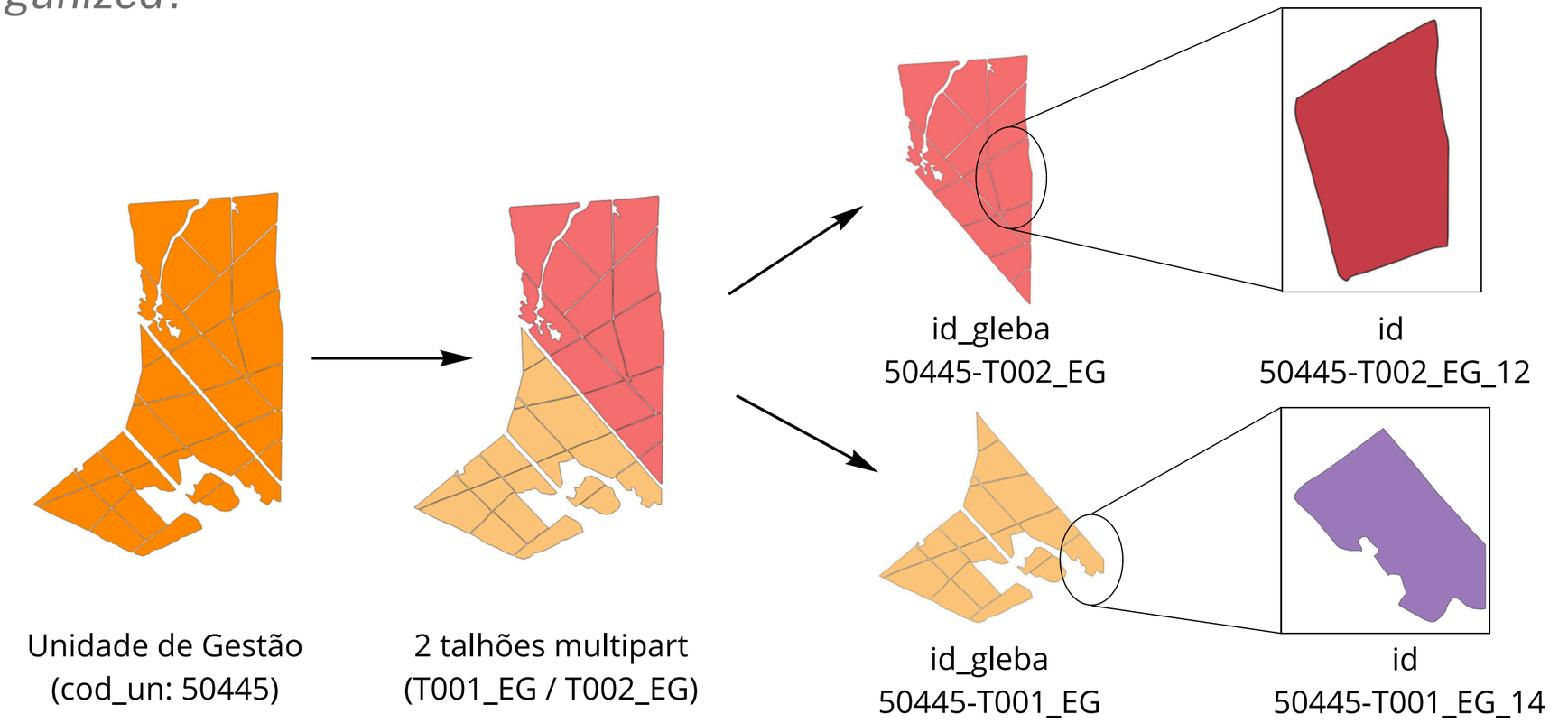
Other pine trees

19 parcels
159 sub-parcels

DATA DESCRIPTION

ORIGINAL DATABASE

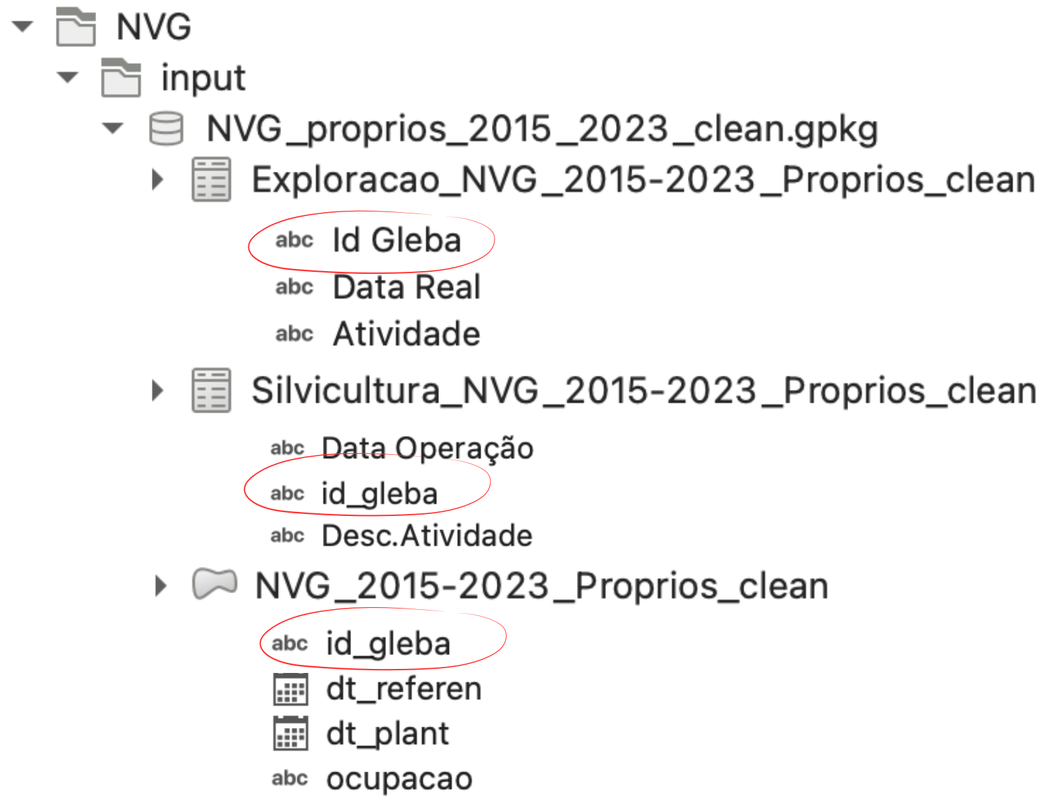
How is is organized?



DATA DESCRIPTION

ORIGINAL DATABASE

How is is organized?



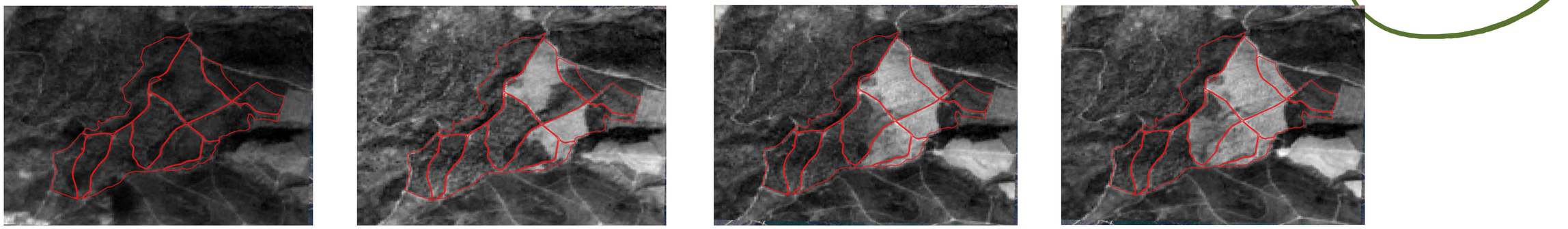
Chronologically sorted database

LIMITAÇÕES DA BASE DE DADOS

ORIGINAL DATABASE

Uncertainty of the clear-cut dates

1st clear-cut date on the Navigator's database



(Tile 29TPE)

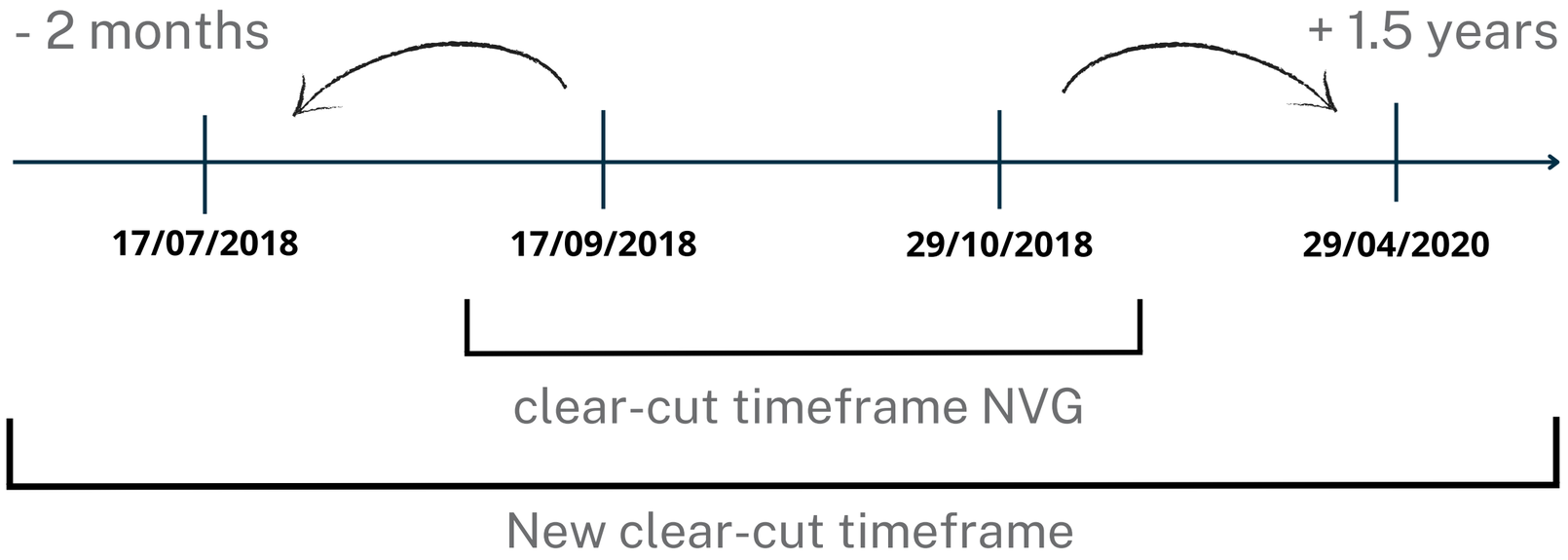
LIMITAÇÕES DA BASE DE DADOS

ORIGINAL DATABASE

Uncertainty of the clear-cut dates

New clear-cut timeframe based on NVG clear-cut dates

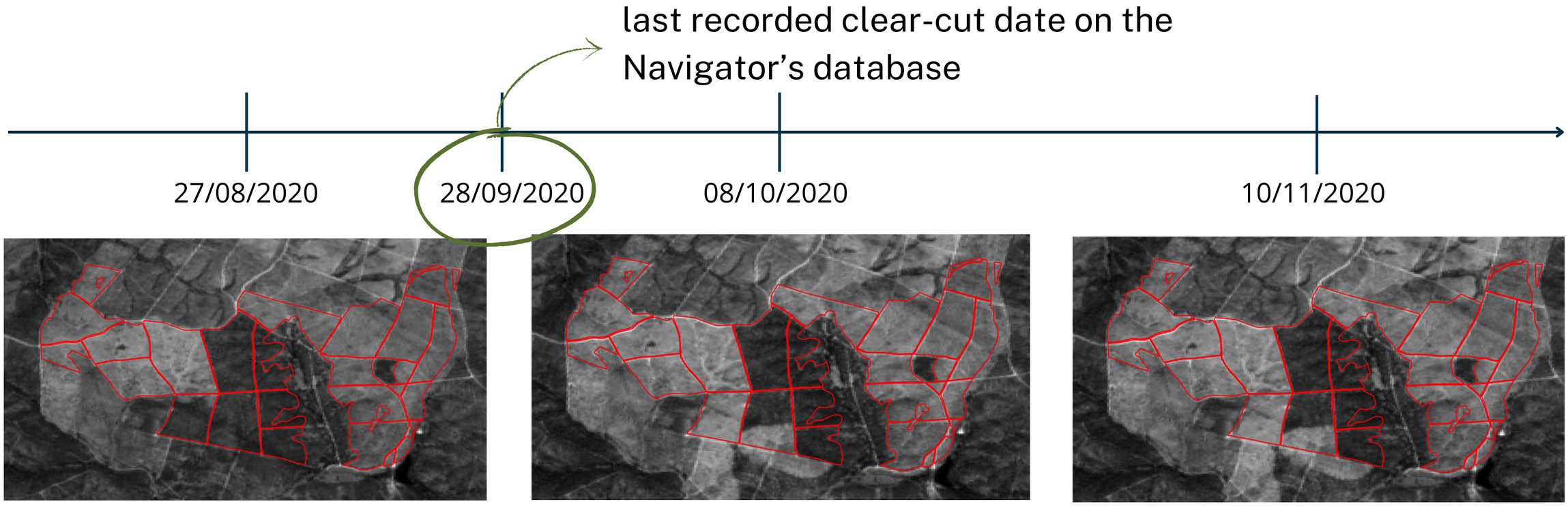
[NVG first date - 2 months; NVG last date + 1.5 years]



LIMITAÇÕES DA BASE DE DADOS

ORIGINAL DATABASE

Heterogeneity of sub-parcel's clear-cuts



(Tile 29TPE)

METHODOLOGY

ESTIMATE CLEAR-CUT DATES AT PIXEL LEVEL

Continuous Change Detection algorithm (CCD)

- Designed to detect land cover changes continuously over time using satellite data

  ↳ *SENTINEL-2 COLLECTION WITH 'S2CLOUDLESS' FILTER*

- Detects time series breaks when the difference between consecutive observations and predicted values surpasses a set threshold.

- The breaks are typically indicative of land cover changes or disturbances (such as wildfires, deforestation)

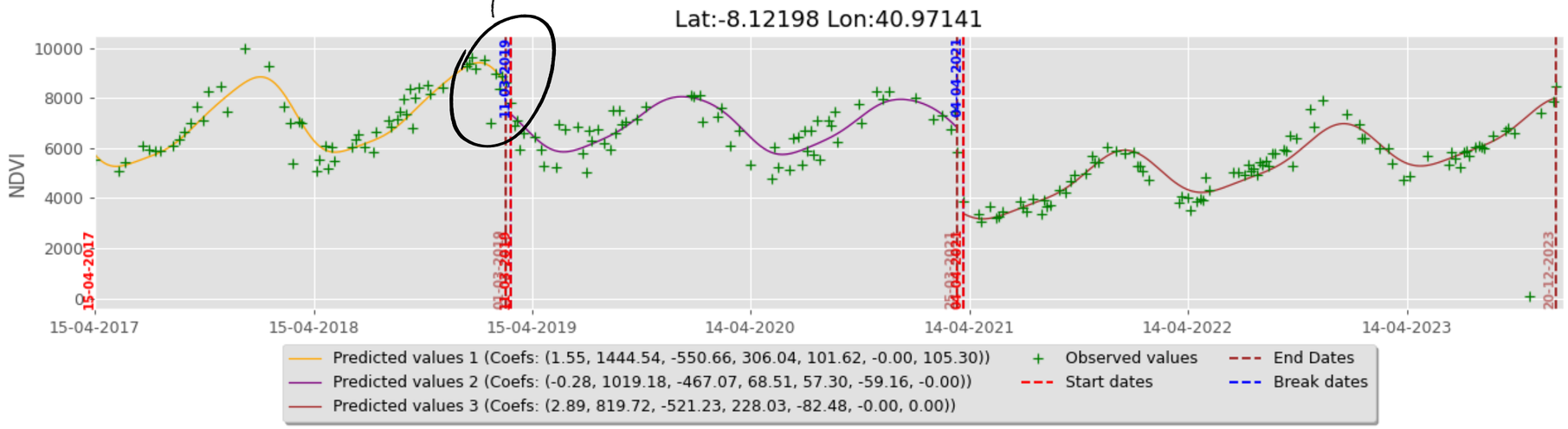
METHODOLOGY

ESTIMATE CLEAR-CUT DATES AT PIXEL LEVEL

Continuous Change Detection algorithm (CCD)

first break date (tBreak)

CCD output for each pixel:

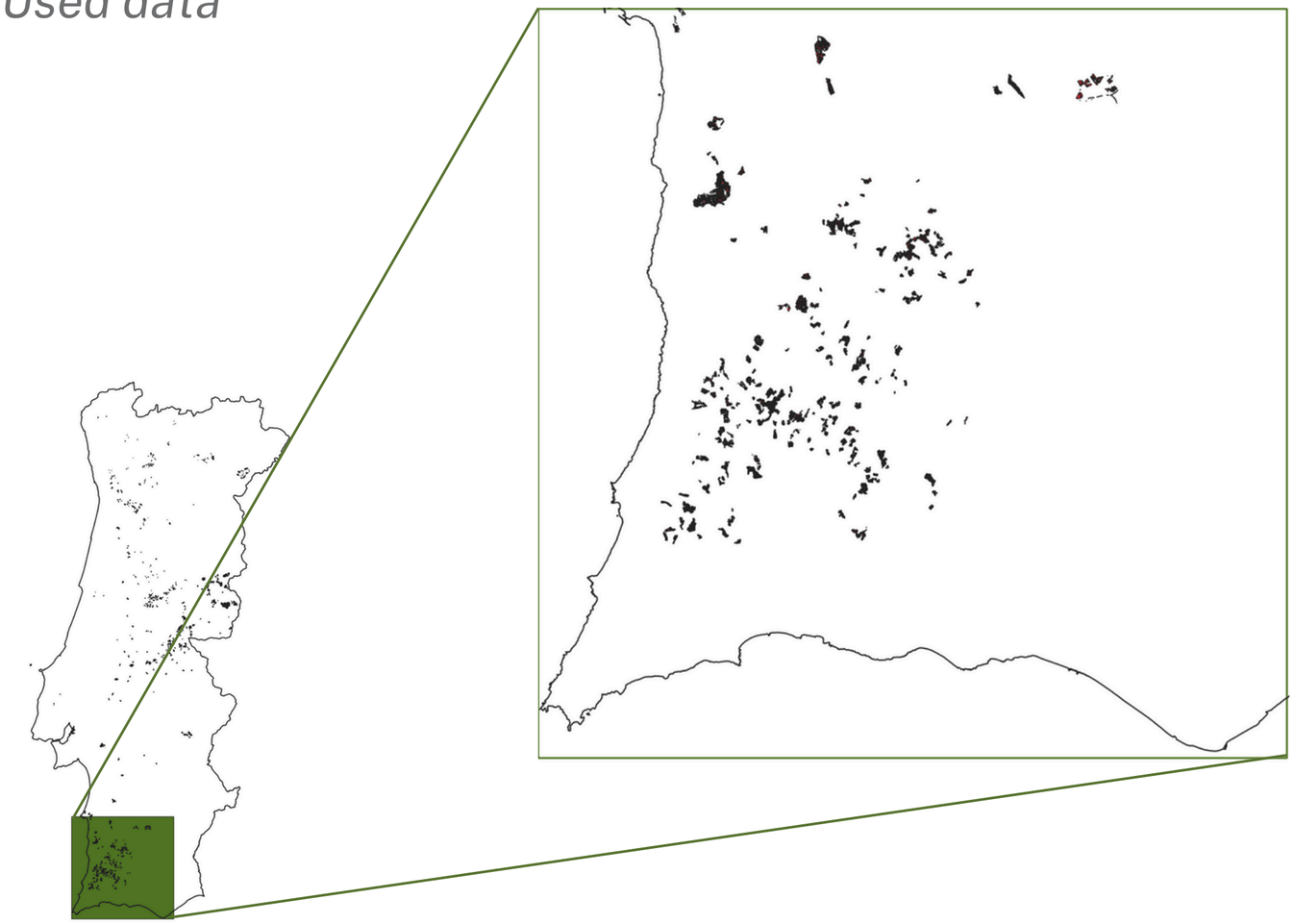


Only the first tBreak occurring within the specified timeframe was taken into account.

METHODOLOGY

ESTIMATE CLEAR-CUT DATES AT PIXEL LEVEL

Used data



TILE 29SNB
251 parcels
3793 sub-parcels

METHODOLOGY

ESTIMATE CLEAR-CUT DATES AT PIXEL LEVEL

Visual Analysis of Parcels

When the CCD results do not match the trend showed on the S2 images → ADD **NEW ATTRIBUTES** TO ATTRIBUTES TABLES

NC (“not cut”) - when the sub-parcel is not cut within the defined timeframe.
(values between 0 and 1 representing an estimation proportion of the sub-parcel that was not cut)

(“Error CCD”) - when the S2 images show a clear-cut but the algorithm identifies a different tBreak date or fails to detect one at all.

- **ECCD1 -** S2 image date immediately before the first observed clear-cut *(date-type)*
- **ECCD2** - S2 image date after the last cut in the sub-parcel was identified *(date-type)*

METHODOLOGY

ESTIMATE CLEAR-CUT DATES AT PIXEL LEVEL

Visual Analysis of Parcels

When the CCD results do not match the trend showed on the S2 images → ADD **NEW ATTRIBUTES** TO ATTRIBUTES TABLES

NC (“not cut”) - when the sub-parcel is not cut within the defined timeframe.
(values between 0 and 1 representing an estimation proportion of the sub-parcel that was not cut)

(“Error CCD”) - when the S2 images show a clear-cut but the algorithm identifies a different tBreak date or fails to detect one at all.

- **ECCD1 -** S2 image date immediately before the first observed clear-cut *(date-type)*
- **ECCD2** - S2 image date after the last cut in the sub-parcel was identified *(date-type)*

tolerância de +- 5 dias

METHODOLOGY

ESTIMATE CLEAR-CUT DATES AT PIXEL LEVEL

Visual Analysis of Parcels

CCD-correct sub-parcels

A sub-parcel is considered correct when:

1. Drop_date is not NULL, and ECCD1, ECCD2, and NC variables are NULL
2. Drop_date is NULL and NC = 1

Isolates NULL's: pixels with a NULL drop_date even though the CCD returns an accurate description of the temporal dynamics of the cuts

CCD-incorrect sub-parcels

A sub-parcel is considered incorrect when it does not meet the conditions above

METHODOLOGY

ESTIMATE CLEAR-CUT DATES AT PIXEL LEVEL

Visual Analysis of Parcels

CCD-correct sub-parcels

A sub-parcel is considered correct when:

1. Drop_date is not NULL, and ECCD1, ECCD2, and NC variables are NULL
2. Drop_date is NULL and NC = 1

READ AT PIXEL LEVEL

Isolates NULL's: pixels with a NULL drop_date even though the CCD returns an accurate description of the temporal dynamics of the cuts

CCD-incorrect sub-parcels

A sub-parcel is considered incorrect when it does not meet the conditions above

READ AT SUB-PARCEL LEVEL

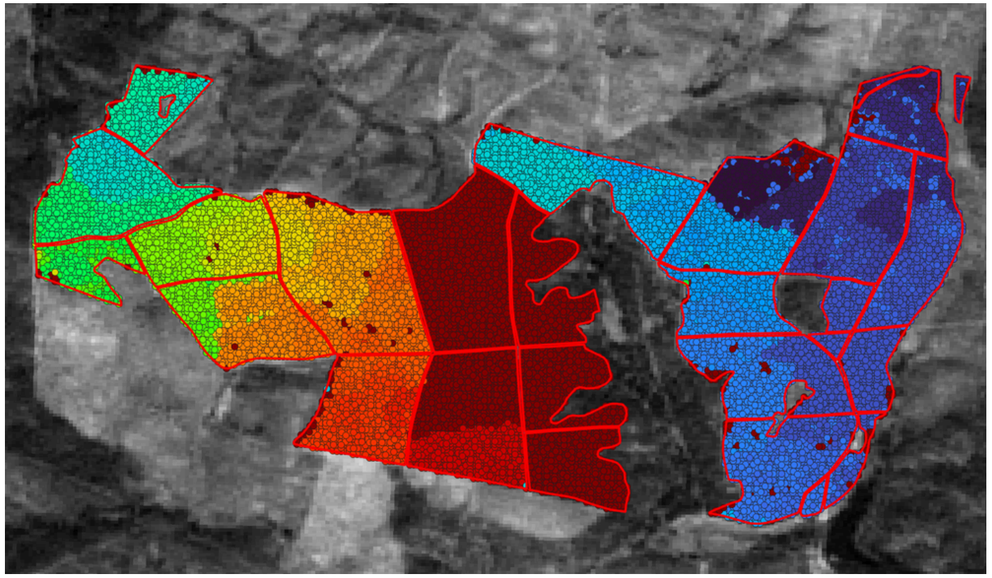
RESULTS

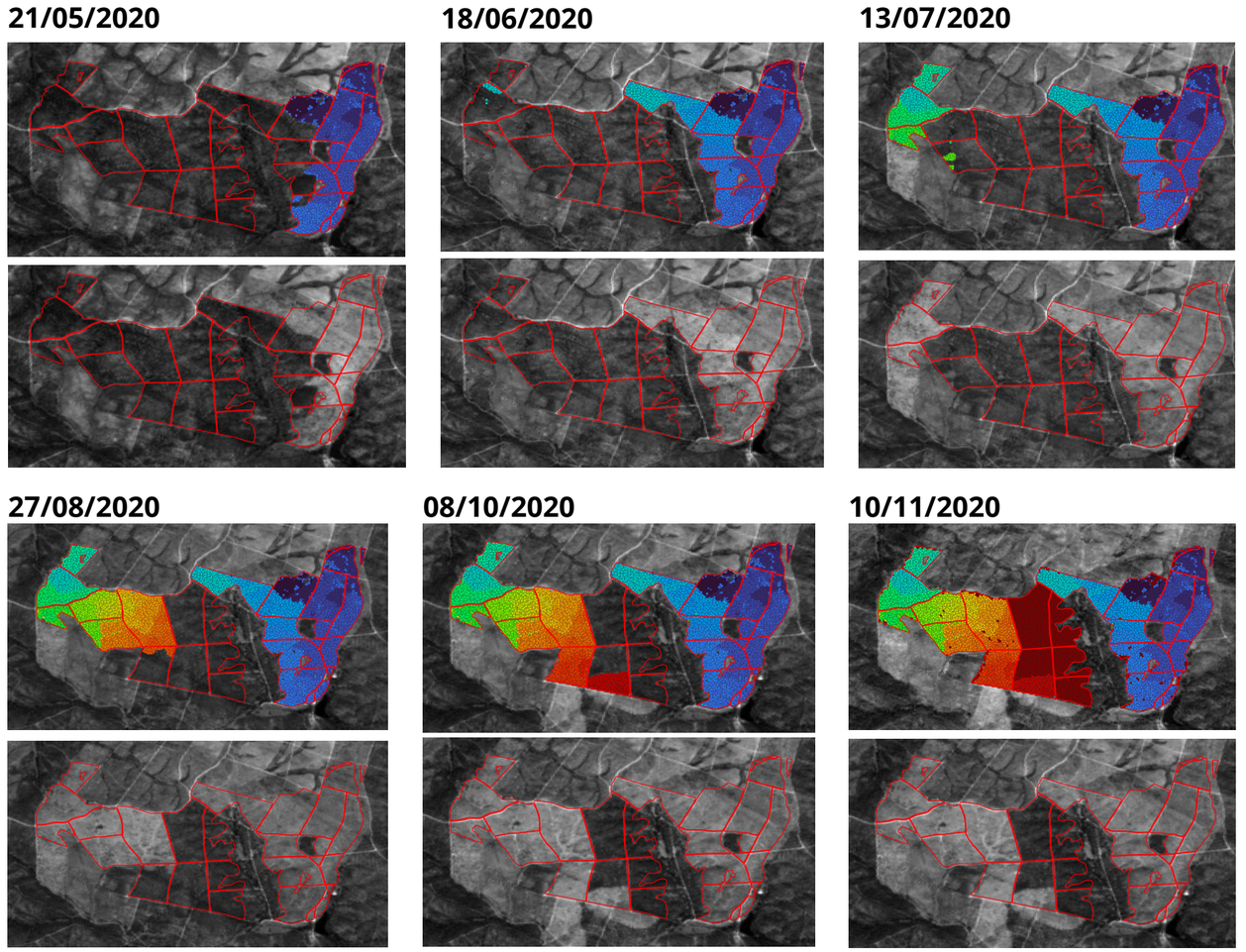
ESTIMATE CLEAR-CUT DATES AT PIXEL LEVEL

Visual Analysis of Parcels

CCD-correct sub-parcels

Id gleba 53001-T006_EG





(Tile 29TPE)

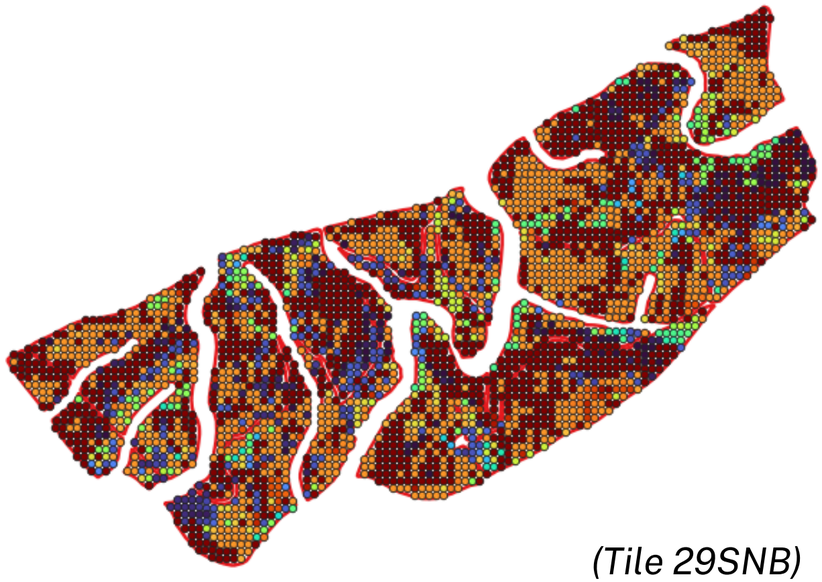
RESULTS

ESTIMATE CLEAR-CUT DATES AT PIXEL LEVEL

Visual Analysis of Parcels

CCD-incorrect sub-parcels

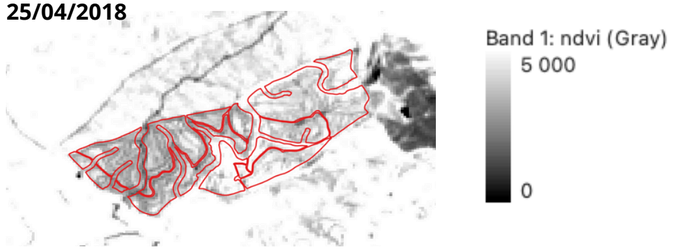
Id gleba 51316-T003_EG



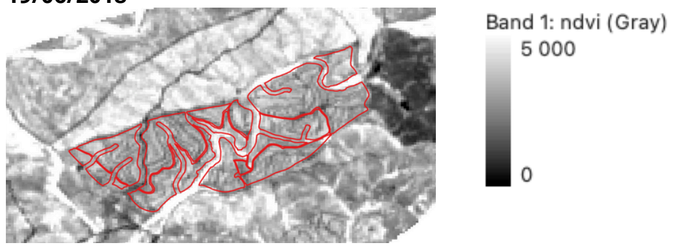
(Tile 29SNB)

26/03/2018

25/04/2018



19/06/2018



id	ECCD1	ECCD2
51316-T003_EG_07	26/03/2018	25/04/2018
51316-T003_EG_06	26/03/2018	25/04/2018
51316-T003_EG_09	26/03/2018	25/04/2018
51316-T003_EG_08	26/03/2018	25/04/2018
51316-T003_EG_11	26/03/2018	19/06/2018
51316-T003_EG_10	26/03/2018	19/06/2018
51316-T003_EG_13	26/03/2018	19/06/2018
51316-T003_EG_12	26/03/2018	19/06/2018
51316-T003_EG_15	26/03/2018	25/04/2018
51316-T003_EG_14	26/03/2018	19/06/2018
51316-T003_EG_17	25/03/2018	19/06/2018
51316-T003_EG_01	26/03/2018	25/04/2018
51316-T003_EG_16	26/03/2018	19/06/2018
51316-T003_EG_03	26/03/2018	25/04/2018
51316-T003_EG_02	26/03/2018	25/04/2018
51316-T003_EG_05	26/03/2018	25/04/2018
51316-T003_EG_04	26/03/2018	25/04/2018

RESULTS

ESTIMATE CLEAR-CUT DATES AT PIXEL LEVEL

Revised Reference Database

Id_gleba ← NVG original database

Id

Start_date

End_date

based on NVG original database

Drop_date ← from CCD results

ECCD1

ECCD2

NC

visual analysis results

Data0

Data1

based on the CCD and visual analysis results

RESULTS

ESTIMATE CLEAR-CUT DATES AT PIXEL LEVEL

Revised Reference Database

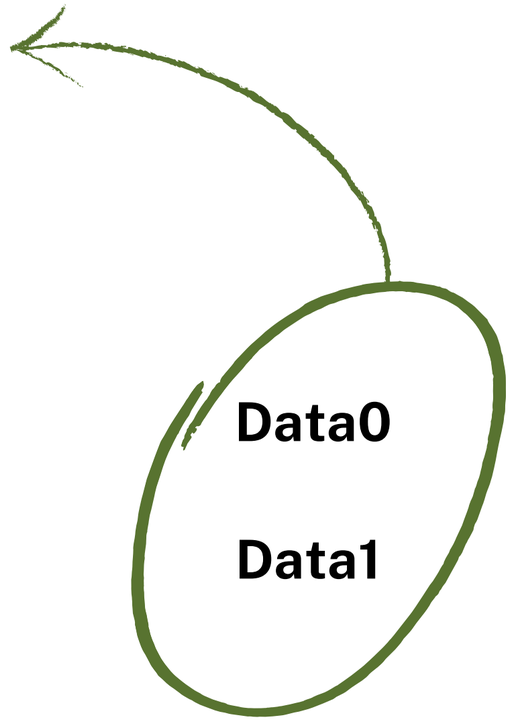
DEPEND ON THE CORRECTNESS OF THE CCD RESULTS

Correct CCD results

↳ based on the drop_date value

Incorrect CCD results

↳ based on the ECCD1, ECCD2 and NC



The ideal difference between data0 and data1 is **10 days**

RESULTS

ESTIMATE CLEAR-CUT DATES AT PIXEL LEVEL

Revised Reference Database - Conditions

Data0

When drop_date is not NULL:

If ECCD1 = 01-01-1000 then data0 = start_date

If ECCD1 = date then data0 = ECCD1

If ECCD1 = NULL AND NC = NULL then data0 = (drop_date – 5 days)

If ECCD1 = NULL AND NC != NULL then data0 = NULL

When drop_date is NULL:

If ECCD1 = NULL AND NC = NULL then data0 = NULL

If ECCD1 is not NULL AND ECCD1 != ’01-01-1000’ then data0 = ECCD1

If ECCD1 = 01-01-1000 then data0 = start_date

Data1

When drop_date is not NULL:

If ECCD2 = 01-01-1000 then data1 = end_date

If ECCD2 = date then data1 = ECCD2

If ECCD2 = NULL AND NC = NULL then data1 = (drop_date – 5 days)

If ECCD2 = NULL AND NC != NULL then data1 = NULL

When drop_date is NULL:

If ECCD2 = NULL AND NC = NULL then data1 = NULL

If ECCD2 is not NULL AND ECCD2 != ’01-01-1000’ then data1 = ECCD2

If ECCD2 = 01-01-1000 then data1 = end_date

RESULTS

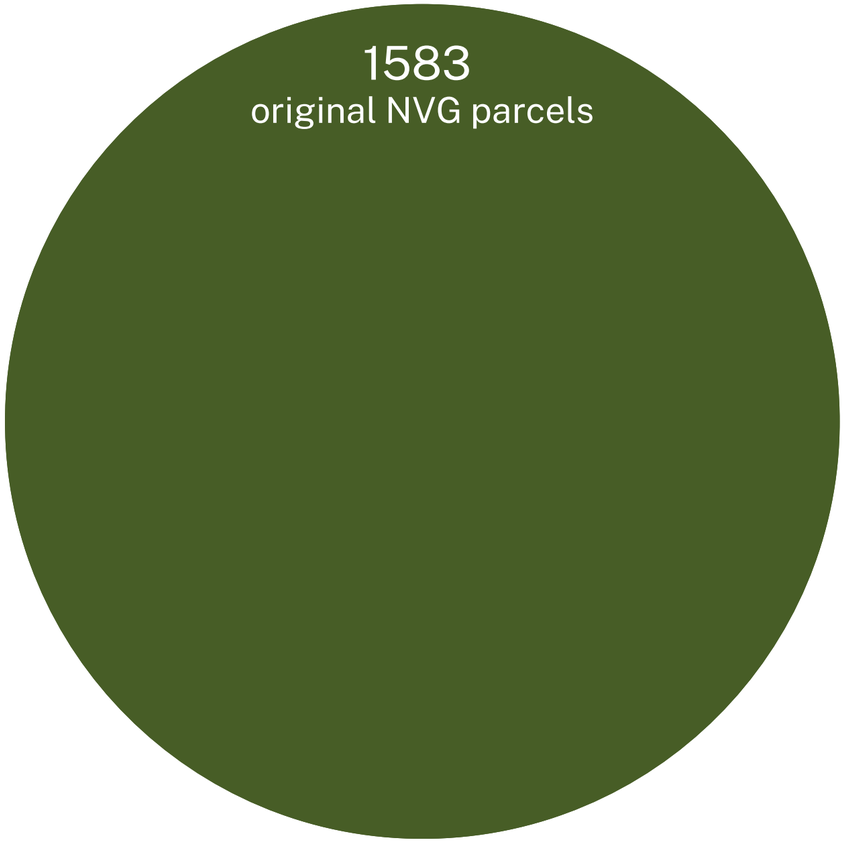
ESTIMATE CLEAR-CUT DATES AT PIXEL LEVEL

Revised Reference Database - Tile 29SNB

RESULTS

ESTIMATE CLEAR-CUT DATES AT PIXEL LEVEL

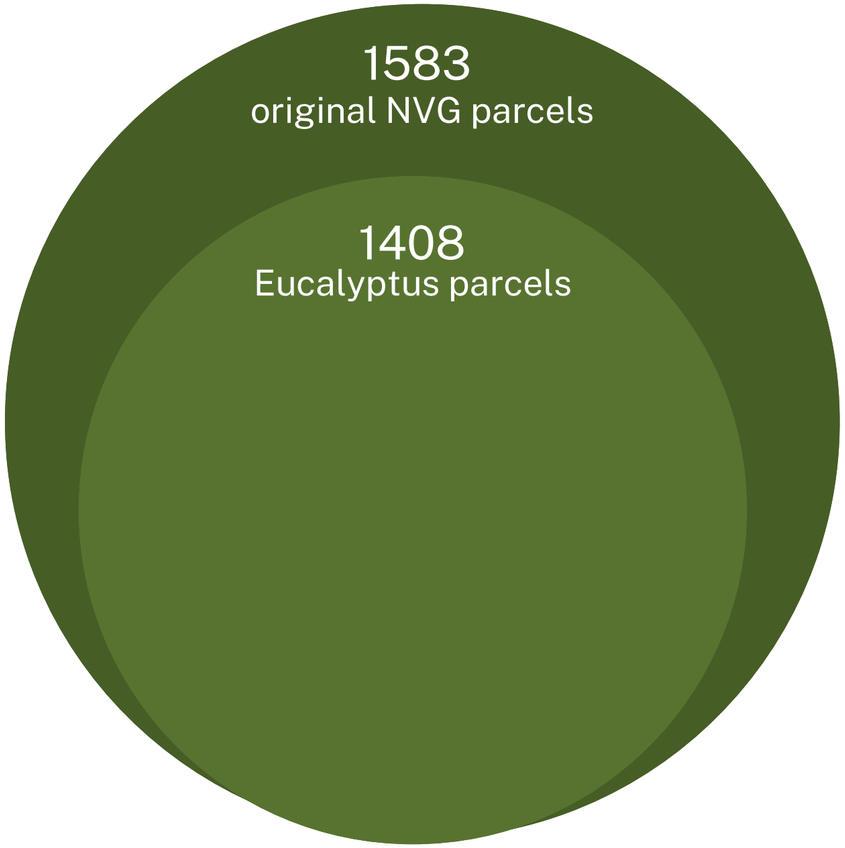
Revised Reference Database - Tile 29SNB



RESULTS

ESTIMATE CLEAR-CUT DATES AT PIXEL LEVEL

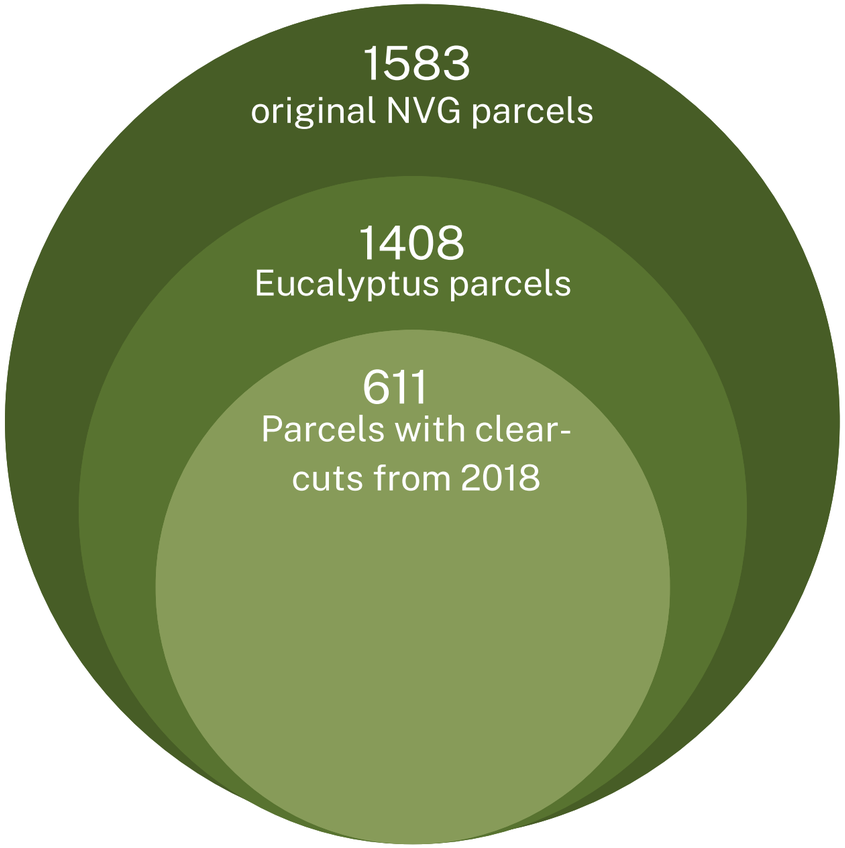
Revised Reference Database - Tile 29SNB



RESULTS

ESTIMATE CLEAR-CUT DATES AT PIXEL LEVEL

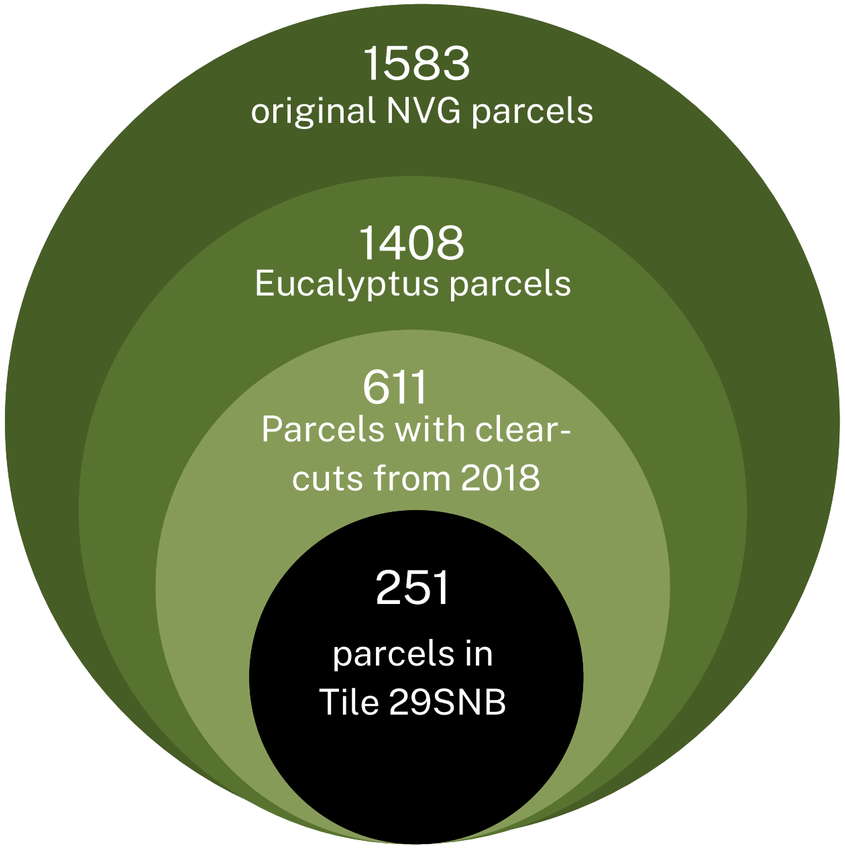
Revised Reference Database - Tile 29SNB



RESULTS

ESTIMATE CLEAR-CUT DATES AT PIXEL LEVEL

Revised Reference Database - Tile 29SNB



RESULTS

ESTIMATE CLEAR-CUT DATES AT PIXEL LEVEL

Revised Reference Database - Tile 29SNB

74.3% CCD-correct sub-parcels

25.7% CCD-incorrect sub-parcels

21.7% sub-parcels updated with new date range interval

*** 20 sub-parcels with no pixels

DISCUSSION

The primary focus was to assign an **estimated clear-cut date to each sub-parcel**

DISCUSSION

The primary focus was to assign an **estimated clear-cut date to each sub-parcel**

A **pixel-level approach** was explored instead

DISCUSSION

The primary focus was to assign an **estimated clear-cut date to each sub-parcel**

A **pixel-level approach** was explored instead

flexible dynamic, **identifying smaller-scale clear-cuts** that might be overlooked in a sub-parcel-based analysis

DISCUSSION

The primary focus was to assign an **estimated clear-cut date to each sub-parcel**

A **pixel-level approach** was explored instead

flexible dynamic, **identifying smaller-scale clear-cuts** that might be overlooked in a sub-parcel-based analysis

The CCD algorithm generally **performed well in identifying cutting activities**, but there was still a **need to revise the results**

DISCUSSION

The primary focus was to assign an **estimated clear-cut date to each sub-parcel**

A **pixel-level approach** was explored instead

flexible dynamic, **identifying smaller-scale clear-cuts** that might be overlooked in a sub-parcel-based analysis

The CCD algorithm generally **performed well in identifying cutting activities**, but there was still a **need to revise the results**

Visual analysis of parcels from tile 29SNB

DISCUSSION

The primary focus was to assign an **estimated clear-cut date to each sub-parcel**

A **pixel-level approach** was explored instead

flexible dynamic, **identifying smaller-scale clear-cuts** that might be overlooked in a sub-parcel-based analysis

The CCD algorithm generally **performed well in identifying cutting activities**, but there was still a **need to revise the results**

Visual analysis of parcels from tile 29SNB

Revised Reference Database

DISCUSSION

CHALLENGES AT USER'S PERSPECTIVE

↳ Sub-parcels with a high count of isolated NULL

↳ CCD-incorrect sub-parcels with a large temporal uncertainty.

DISCUSSION

CHALLENGES AT USER'S PERSPECTIVE

↳ Sub-parcels with a high count of isolated NULL

- introduces noise into the database
- data0 and data1 will be NULL

↳ CCD-incorrect sub-parcels with a large temporal uncertainty.

estimate a clear-cut date using the most frequent pixel value within the sub-parcel

DISCUSSION

CHALLENGES AT USER'S PERSPECTIVE

↳ Sub-parcels with a high count of isolated NULL

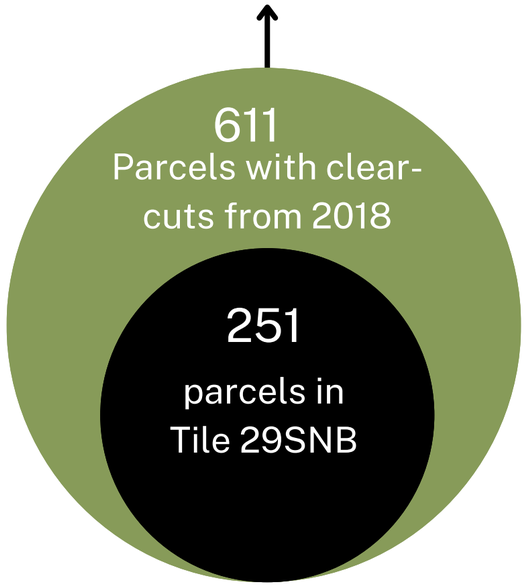
↳ CCD-incorrect sub-parcels with a large temporal uncertainty.

- large time intervals between data0 and data1
- less precise data0 and data1 values

EXTENSIONS OF THE RESULTS TO ALL CONTINENTAL PORTUGAL

ESTIMATE CLEAR-CUT DATES AT PIXEL LEVEL

Revised Reference Database - All tiles



8018 sub-parcels → *** → 7980 sub-parcels

- 72.7% CCD-correct sub-parcels
- 27.3% CCD-incorrect sub-parcels
- 24.3% sub-parcels updated with new date range interval

*** 38 sub-parcels with no pixels

CONCLUSION

- First comprehensive reference database
  - ↳ polygon format
  - ↳ covers Continental Portugal
  - ↳ clear-cuts
- Semi-automatic method
  - ↳ can be extended to other reference database (e.g. areas ardidas do ICNF)
  - ↳ feasible with limited resources
- Development of machine learning change detection algorithms
  - ↳ train
  - ↳ validation

Thank you!

Inês Silveira