

Reunião DGT/INCD/LIP, 4 de julho de 2024

Breve descrição da colaboração do problema de criação de um produto de perda recente de floresta e mato com base em análise automática de imagens de satélite

Manuel Campagnolo, ISA, CEF, ULisboa

Objectivo da colaboração ISA/DGT:

- Desenvolver metodologias eficazes à escala nacional e eficientes a nível computacional para a criação sistemática de um produto nacional
- Produto em formato vetorial de delimitação de manchas superiores a 0.5 ha de perda recente de floresta e mato
- A geração do produto é baseada em análise automática de imagens de satélite

Aspetos em desenvolvimento:

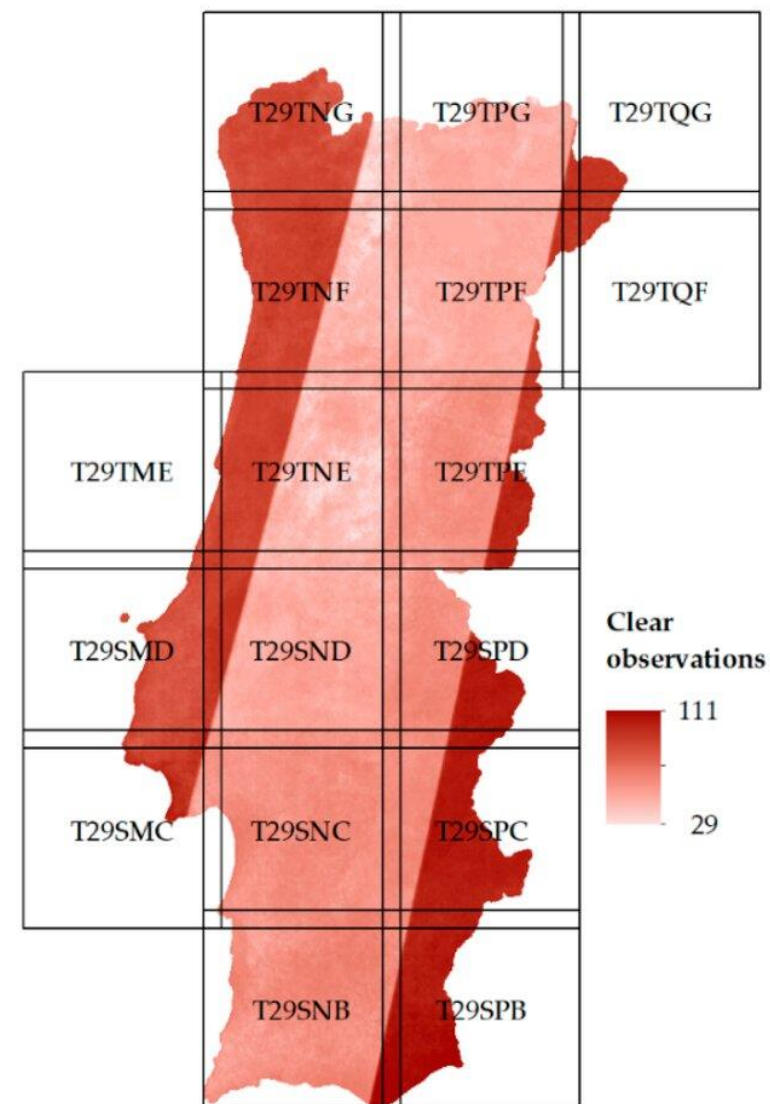
- Dados de referência
- Algoritmos de deteção de alterações e classificação
- Validação dos resultados (precisão do algoritmo)
- Implementação

Dados

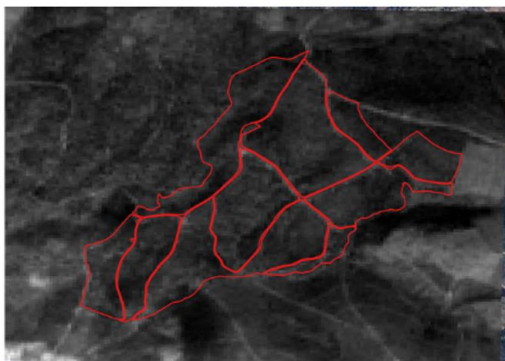
Imagens de satélite

- Coleção de imagens Sentinel-2 para Portugal Continental
- Cada tile 100 km por 100 km, 13 bandas, ocupa ~1Gb
- A resolução temporal é de 5 dias
- As imagens Sentinel-2 estão disponíveis desde 2016

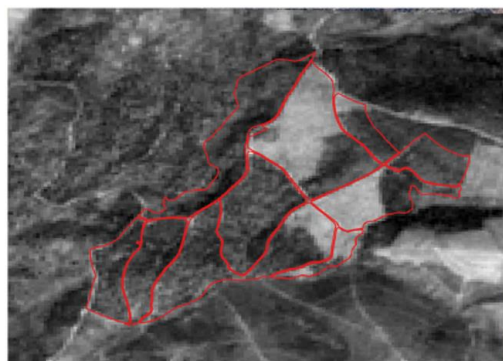
Dados de referência



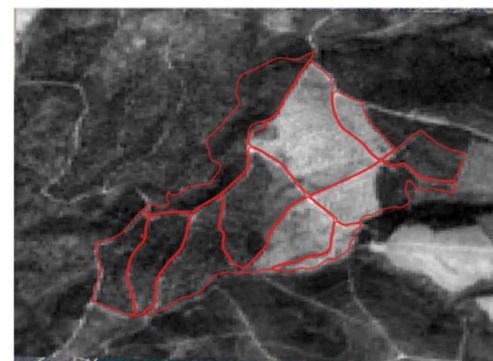
Dados: informação vetorial sobre cortes



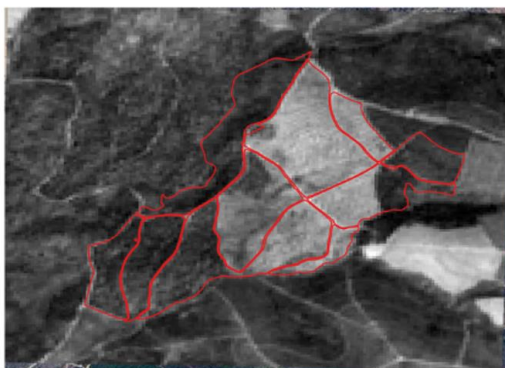
05/04/2018



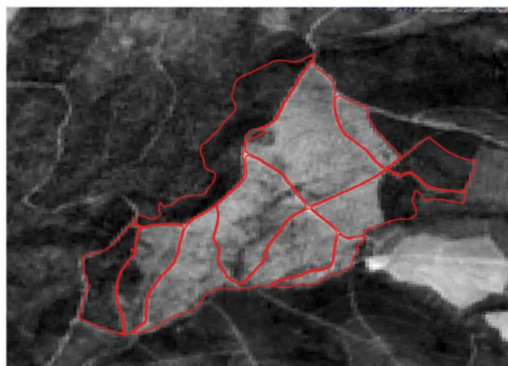
17/04/2018



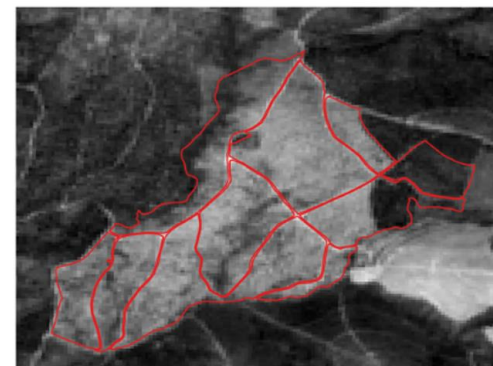
25/04/2018



27/04/2018



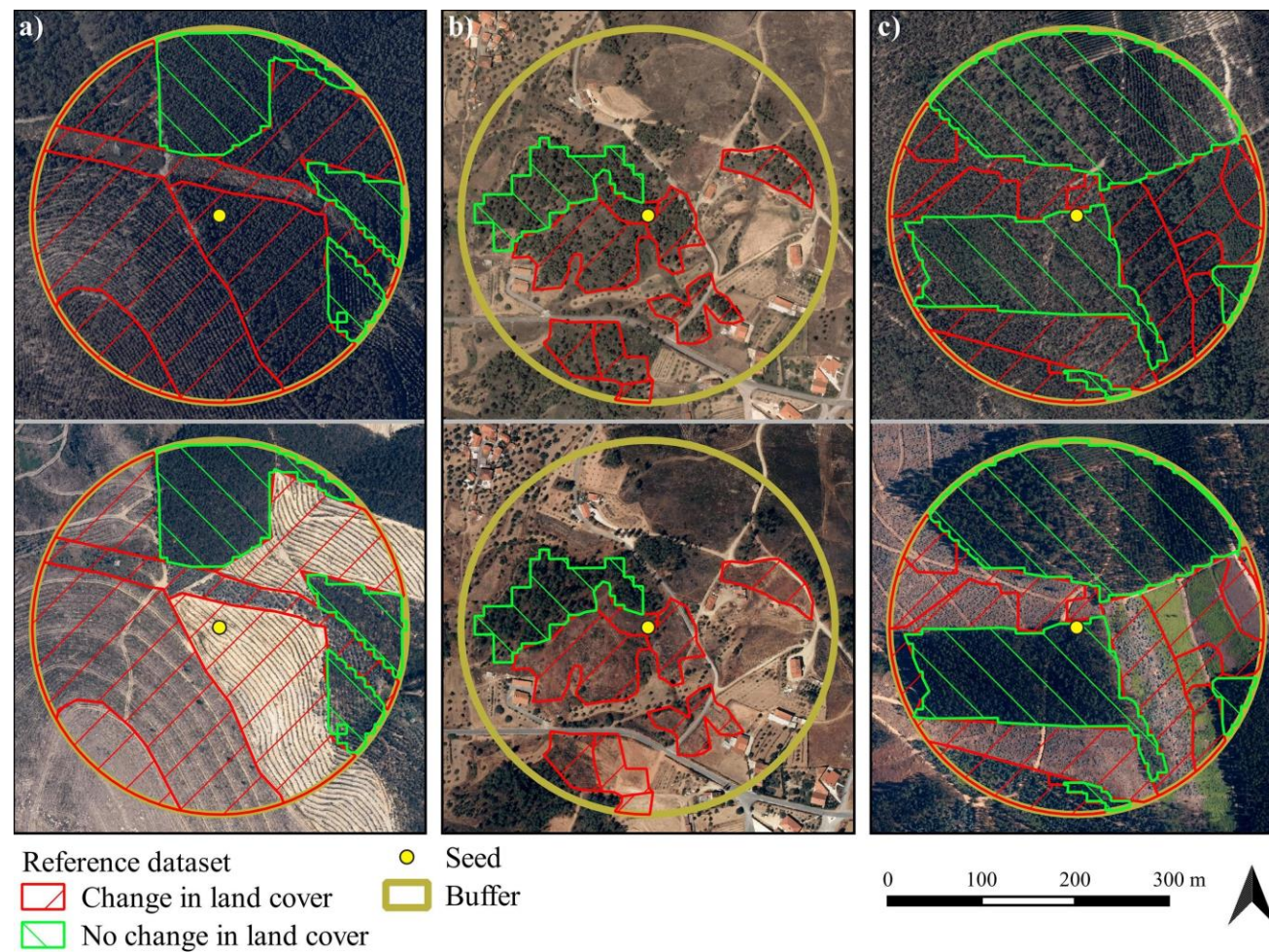
05/05/2018



15/05/2018

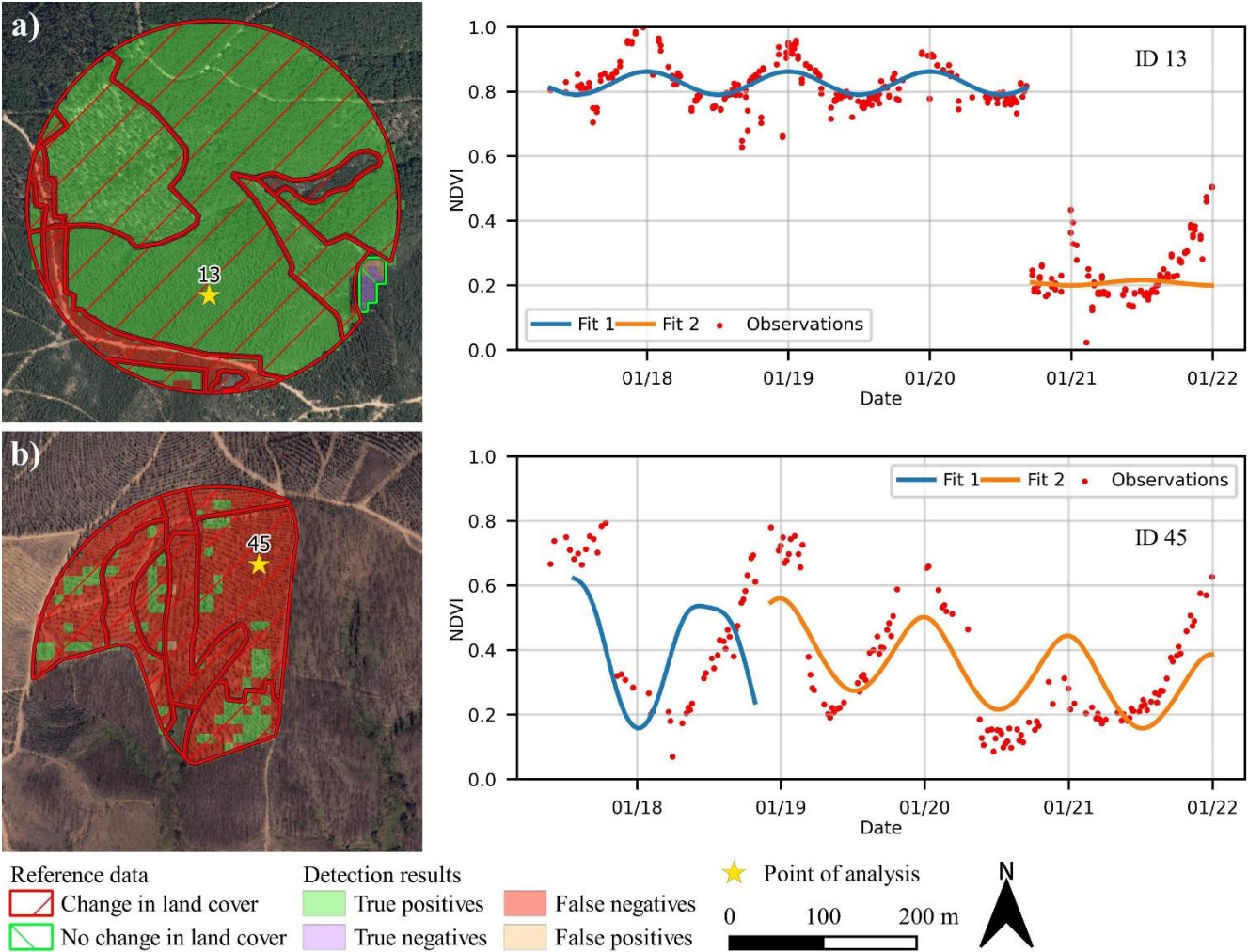
Série temporal do índice espectral NDVI derivado de imagens Sentinel-2 (resolução 10m) sobre um talhão no qual foram realizados cortes rasos de eucalipto

Dados de referência



Exemplo de dados de referência para treinar um modelo de detecção (CCD)

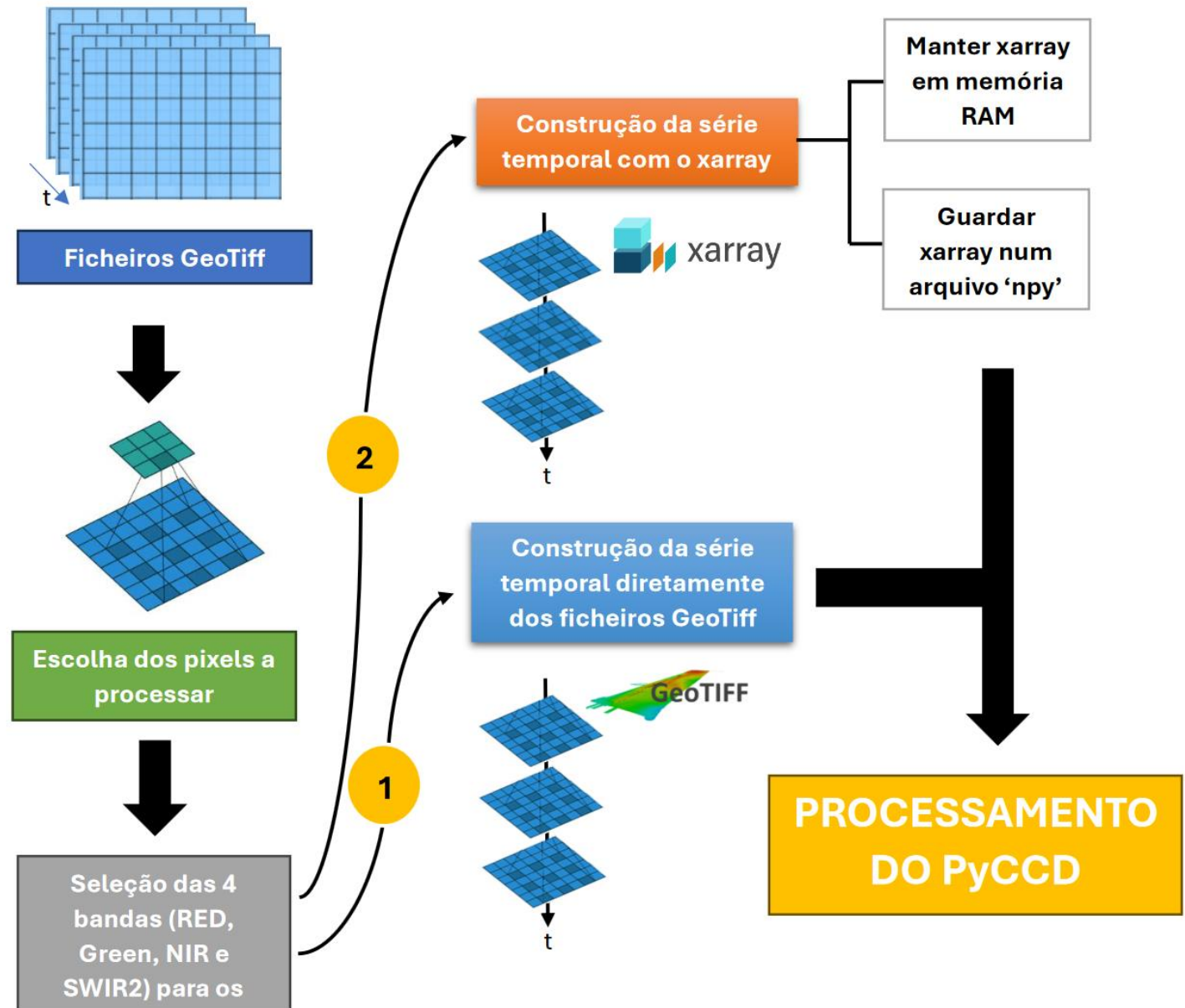
Processamento: algoritmo de detecção de alterações (CCD) aplicado a cada pixel



Ajustamento de segmentos temporais usando CCD

Étapas de processamento

- Upload em tempo real das imagens da coleção Sentinel-2, com correção atmosférica
- Upload na mascara de nuvens (opcional)
- Leitura da série temporal por pixel da imagem
- Processamento da série temporal para cada pixel (processo a paralelizar)
- Criação dos mapas de datas de alterações



Estimativas de recursos computacionais para criação de uma cobertura nacional do produto (uma data)

Imagens préprocessadas (apenas zonas de floresta e mato; 3 bandas espectrais)

- $120 * 10^6$ pixels; 8 anos de imagens; corresponde no total a ~1Tb
- Leitura da série temporal por pixel da imagem: o tempo de leitura depende do formato dos ficheiros (tif; npy; ...)
- Processamento das séries temporais: estimativa de 34560 core-hours para a cobertura nacional (seriam 270 h de processamento para 128 cores). Parte do código (regressão LASSO) está a ser re-escrito em Cython, com uma estimativa de 15x aceleração relativamente aos valores acima