Plan Overview

A Data Management Plan created using DMPonline

Title: Assessing fire-induced vegetation structure changes by integrating spaceborne LiDAR and multi-frequency satellite data

Creator: Martín Domínguez Durán

Principal Investigator: Johannes Reiche, Martín Domínguez Durán, Guido van der Werf

Data Manager: Martín Domínguez Durán

Project Administrator: Martín Domínguez Durán

Affiliation: Wageningen University and Research (Netherlands)

Funder: European Commission

Template: Data Management Plan | Wageningen University and Research

ORCID iD: 0000-0002-4327-4349

ORCID iD: 0009-0007-0134-9588

ORCID iD: 0000-0001-9042-8630

Project abstract:

Fire is a key driver of ecosystem change and carbon emissions globally, yet our ability to quantify its structural impacts on vegetation remains limited. While remote sensing has improved fire detection and burn severity mapping, accurately assessing fire-induced changes in vegetation structure continues to pose a major challenge—primarily due to the scarcity of consistent, high-quality reference data. This research aims to bridge critical knowledge gaps by leveraging spaceborne LiDAR data and integrating it with multi-frequency satellite datasets, including optical and synthetic aperture radar (SAR) imagery. The project is structured around three main objectives: (1) facilitating the analysis of vegetation structure change using comparable spaceborne LiDAR measurements (2) spatially characterize the impacts of fire in the vegetation structure of two distinct fire-prone eco-regions by integrating LiDAR measurements with multi-frequency satellite data; and (3) quantifying fire-induced structural changes globally across diverse ecosystems. This study seeks to advance our understanding of post-fire vegetation dynamics and reduce uncertainties in biomass loss and fire emission estimates. The findings will not only contribute to vegetation monitoring and Earth system science but also support climate mitigation efforts by improving the representation of fire impacts in carbon cycle models.

ID: 178121

Start date: 01-02-2025

End date: 01-05-2029

Last modified: 20-05-2025

Copyright information:

Assessing fire-induced vegetation structure changes by integrating spaceborne LiDAR and multi-frequency satellite data

A. Describe the research project

1. Name researcher (please, add your full name):

Martín Domínguez Durán

2. What is the name of your department(s)?

- Environmental Sciences

3. What is the name of your chair group(s) or business unit(s)? English name and abbreviation for chair groups from this page; business units from this page (expand to Wageningen Research and keep expanding to find your specific division / group). Examples: Bioprocess Engineering (BPE) or Contract Research Organization (CRO).

Meteorology and Air Quality (MAQ)
Laboratory of Geo-Information and Remote Sensing (GRS)

4. Describe the organisational context of your research project.

DMP version (or date last modified)	1 (05/2025)
Supervisor / (co-)promotors	Guido van der Werf, Johannes Reiche
Graduate School (WU only)	PE&RC
Start date of project	01/02/2025
End date of project	01/05/2025
Project number	
Funding body	European Union

5. Give a short description of your research project.

Title	Assessing fire-induced vegetation structure changes by integrating spaceborne LiDAR and multi-frequency satellite data
Summary	he research project aims to assess vegetation structure changes caused by fire by integrating spaceborne LiDAR data (GEDI, ICESat-2) with multi-frequency satellite observations (Sentinel-1/2, NISAR, BIOMASS). The main objectives are to develop a framework for consistent LiDAR data acquisition, evaluate fire impacts in the Iberian Peninsula, and model global vegetation changes to improve fire emission estimates. The project uses existing public and national datasets, generates derived vegetation and fire impact metrics, and makes most outputs openly available to support researchers, environmental agencies, and climate modelers.

6. List the individuals responsible for the following data management tasks.

Data collection	Martín Domínguez Durán
Data quality	Martín Domínguez Durán - Johannes Reiche
Storage and backup	Martín Domínguez Durán - Aldo Bergsma
Data archiving / publishing	Martín Domínguez Durán - Guido van der Werf
Data stewardship / support	Braden Owsley

7. I have requested a review of this data management plan from:

- No review requested.

8. Name of the data management support staff and / or data steward consulted during the preparation of this plan and date of consultation.

Braden Owsley (GRS)
Aldo Bergsma (GRS)

B. Describe the data to be collected, software used, file formats and data size.

9. Will you use existing data for this project?

- Yes. Please specify below which data (e.g. DOI, URL, or storage location) and the terms of use (e.g. licence).
- **GEDI (Global Ecosystem Dynamics Investigation)**
  - **Source**: NASA LP DAAC (https://lpdaac.usgs.gov/)
  - **Format**: HDF5

    - **Terms of Use**: Publicly available under NASA's open data policy.
- **ICESat-2 ATL data**
    - **Source**: NASA NSIDC (https://nsidc.org/data/icesat-2)
    - **Format**: HDF5
    - **Terms of Use**: Publicly available under NASA's open data policy.
- **Sentinel-1 and Sentinel-2**
    - **Source**: Copernicus Open Access Hub (https://scihub.copernicus.eu/)
    - **Format**: GeoTIFF
    - **Terms of Use**: Free and open access under the Copernicus Data Policy.
- **MODIS (Moderate Resolution Imaging Spectroradiometer)**
    - **Source**: NASA LP DAAC (https://modis.gsfc.nasa.gov/data/)
    - **Format**: GeoTIFF
    - **Terms of Use**: Publicly available under NASA’s open data policy.
- **National LiDAR datasets**
    - **Portugal and Spain (ICNF and PNOA)**
        - **Source**: National providers (e.g., Instituto da Conservação da Natureza e das Florestas)
        - **Format**: LAZ
        - **Terms of Use**: May be subject to licensing restrictions; derived, non-sensitive outputs will be shared publicly.

10. Will new data be produced?

- Yes.

11. Please describe the data you expect to generate and / or use in the table below. Include reused existing data as well (as these are files that you manage and store).

File contents	Data type	Software	(Open) file format	Estimated size of each file (range)	Estimated number of files (range)
GEDI/ICESat-2 pairs matched	Geospatial point vector	Python, Pandas, GeoPandas	GeoParquet	50–800 MB	(50-100)
Vegetation structure change metrics (e.g., ΔRH95, ΔAGBD)	Processed geospatial point data	Python, Pandas, GeoPandas	GeoParquet	50–800 MB	(50-100)
Modeled structure change – Iberian Peninsula (raster)	Raster grids	Python, GDAL, Earth Engine	GeoTIFF	500 MB–2 GB	10–50
Modeled structure change – Global (raster)	Raster grids	Python, GDAL, Earth Engine	GeoTIFF	2 GB–15 GB	10–50
Open-source processing code	Python scripts & notebooks	Python	.py, .ipynb	<1MB	50-100

12. Estimate how much data storage you require in total (e.g. by using the information in the table at question 11).

- >1000 GB

C. Storage of data and data documentation / metadata during research

13. Where will the data,code and accompanying documentation / metadata be stored and backed up during the research project (see the WUR Data Storage Finder)? Include platforms you use to share data, collect data on, or send data to for processing or analysis.

- WUR OneDrive for Business - only when an up to date version of the research data is also safely stored on the W:drive or Yoda.
- Other, please specify below the storage medium / system and describe back-up frequency, access management, and geographic location (e.g. within or outside the EU).

During the research project, data and accompanying documentation or metadata will be stored and backed up using a combination of WUR-supported and external platforms, following the recommendations of the WUR Data Storage Finder. Day-to-day files and collaborative work will be stored on WUR OneDrive, which provides secure, access-controlled storage with automatic backup. GitHub (under the user account *mdominguezd*) will be used to share open-source code, processing workflows, and metadata documentation in a version-controlled environment. Final datasets and processed outputs—such as vegetation change layers and derived metrics—will be archived and published on Zenodo, where they will be assigned DOIs for long-term preservation and open access. Large Earth observation datasets from Sentinel, MODIS, GEDI, and ICESat-2 will be accessed and processed via Google Earth Engine (GEE), minimizing local storage needs and allowing scalable

computation in the cloud. If needed, sensitive or restricted datasets—such as national LiDAR reference data—will be stored in GRS NAS, which complies with FAIR data principles and institutional data policies.

D. Structuring your data and information

14. Give a (visual) representation of the folder structure you intend to use.

```
project-root/
│
├── data/
│   ├── raw/
│   │   ├── gedi/
│   │   ├── icesat2/
│   │   ├── sentinel/
│   │   ├── modis/
│   │   ├── lidar_national/
│   │   └── metadata/
│   │
│   ├── intermediate/
│   │   ├── spatial_pairs/
│   │   ├── vegetation_metrics/
│   │   └── fire_masks/
│   │
│   └── processed/
│       ├── iberian_peninsula/
│       └── global/
│
├── outputs/
│   ├── figures/
│   ├── maps/
│   ├── tables/
│   └── final_datasets/
│
├── notebooks/
│   ├── 01_preprocessing.ipynb
│   ├── 02_pair_matching.ipynb
│   ├── 03_analysis.ipynb
│   └── 04_modeling.ipynb
│
├── scripts/
│   ├── utils/
│   ├── download_scripts/
│   └── processing/
│
├── docs/
│   ├── README.md
│   ├── data_management_plan.md
│   ├── metadata_templates/
│   └── publications/
```

```
│
├── environment/
│   ├── requirements.txt
│   └── environment.yml
│
└── LICENSE
```

15. Describe the file naming conventions you intend to use. Please give one or multiple example(s).

Names will be structured using lowercase letters, underscores to separate elements, and version numbers where appropriate. Each file name will include key metadata elements such as region, data type, date (in YYYYMMDD format), and processing level or version.
<region>_<datatype>_<sensor/source>_<YYYYMMDD>_<processing-level>.ext Where:

- region: e.g., *iberian*, *global*, *amazon*
- datatype: e.g., *vegmetrics*, *firemask*, *pairdata*, *modeloutput*
- sensor/source: e.g., *gedi*, *icesat2*, *s1*, *s2*, *modis*
- YYYYMMDD: date of data acquisition or generation
- processing-level: e.g., *raw*, *processed*, *v01*, *final*
- ext: appropriate file extension (.csv, .parquet, .geojson, .tif, etc.)

16. How will you distinguish between versions of files (multiple answers possible)?

- The designation 'vRAW' is added to file names that contain raw unaltered data (before any processing and cleaning). Any alteration of RAW data is done on a copy of the RAW data and appended with a version number which increases with each file modification (e.g. v01, v02, v03 etc.).
- We will use Git versioning for code / scripts.

E. Data documentation and data quality

17. Describe below what data documentation and metadata will accompany the data to help make the data findable, understandable, and reproducible.

- The WUR readme file template (see template at https://doi.org/10.5281/zenodo.7701727).

18. Describe what data and analysis quality controls will be used?

- Supervisors or peers will review the data and results for any anomalies (e.g. unexpected inconsistencies, outliers, correct labeling of data and / or treatments, correct and consistent coding applied, etc.).

F. Working with sensitive data (personal data, ethics), data ownership, sharing and access

19. Who is the (rights)holder of the data (commonly known as the owner of the data)?

- WUR is the (rights)holder of the data.

20. What is the data classification for your project (for example as specified in SmartPIA) taking into account the (privacy) sensitivity of the data?

- Negligible.

21. Is this project registered in SmartPIA?

- No. Please register in SmartPIA in the case (privacy) sensitive data is collected (when applicable: via your supervisor, the project manager, see guidance).

22. Please specify the (sensitive) data and privacy protection measures. Note that any measures undertaken should be consulted with the Information Security Officer (ISO) and Privacy Officer (PO).

- Data is classified as negligible and standard WUR security measures are undertaken.

23. Are there other ethical issues that need to be taken into account which may include approval from ethical committees?

- No.

24. Will there be any intellectual property (IP) rights or alternative applications or routes to impact (such as commercial interests) associated with the data?

- No.

G. Data archiving and publishing

25. Are there reasons to restrict access to the data or limit which data will be made publicly available?

- No.

26. Describe what data from question 11 will be archived internally (e.g. WUR network drive / Yoda@WUR) and not published, for a minimum of 10 years? Include the exact name for the storage medium chosen (see the WUR Data Storage Finder).

- Not applicable as data will be published.

27. What data will be published and made available for reuse via a data repository?

- Data underlying publications or reports. Please specify below which data listed in question 11.

All datasets specified on question 11

28. When will the data be available for reuse, and for how long will the data be available?

- Data will be available for at least 10 years as soon as the article or report is published and not required for any other article publication.

29. Which data repository do you intend to use to make the data findable and accessible (see the WUR Repository Finder)?

- Zenodo.

30. Which metadata standard will be used to describe the data during internal archiving and / or depositing in a data repository?

- Metadata standard from DANS Data Stations, 4TU.ResearchData and / or Zenodo (which often are the DublinCore or DataCite standard).

31. Which licence/terms of use will be applied to the data?

- Open access (Creative Commons Attribution licence (CC BY); anyone can access and reuse with attribution).

H. Data management costs

33. What resources (in time and / or money) will be dedicated to data management, data archiving or publication, and ensuring that data is reusable? Indicate as well how these costs will be covered.

- All costs for 10 year data storage and access management to that data after journal publication or report are covered by the research group / project.