



HAL
open science

Natural grasslands across mainland France: A dataset including a 10 m raster and ground reference points

Léa Panhelleux, Sébastien Rapinel, Laurence Hubert-Moy

► To cite this version:

Léa Panhelleux, Sébastien Rapinel, Laurence Hubert-Moy. Natural grasslands across mainland France: A dataset including a 10 m raster and ground reference points. *Data in Brief*, 2023, pp.109348. 10.1016/j.dib.2023.109348 . halshs-04146201

HAL Id: halshs-04146201

<https://shs.hal.science/halshs-04146201v1>

Submitted on 22 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Data Article

Natural grasslands across mainland France: A dataset including a 10 m raster and ground reference points



Léa Panhelleux, Sébastien Rapinel, Laurence Hubert-Moy*

University of Rennes 2, LETG UMR 6554 CNRS, place du Recteur Henri Le Moal, 35000 Rennes, France

ARTICLE INFO

Article history:

Received 5 May 2023

Revised 20 June 2023

Accepted 22 June 2023

Available online 27 June 2023

Dataset link: [Natural grasslands across mainland France: a dataset including a 10 m raster and ground reference points \(Original data\)](#)

Keywords:

Agriculture

Biodiversity

Field data

Green corridors

LULC

Meadows

Natural habitats

Sentinel-2

ABSTRACT

The data provided here include the first 10 m raster of natural grasslands across mainland France and related ground reference points. The latter consist of 1770 field observations that describe natural and artificial grasslands from respectively a compilation of hundreds of field-based vegetation maps and the European Union Land Parcel Identification System (LPIS). Based on analysis of aerial images, ground reference points were manually extracted from grassland polygons of the field-based vegetation maps and the LPIS within herbaceous areas larger than 30×30 m. The raster data of natural grasslands were derived from five annual 10 m land cover maps of France from 2016–2020. Pixels classified as “grassland” every year from 2016–2020 were considered natural grasslands, while those classified as “crop” at least once were considered artificial grasslands. Validation using the ground reference points revealed that natural and artificial grasslands were accurately mapped (overall accuracy = 86%). The ground reference points, publicly available in GeoJSON vector format, can be used as training or test samples for spatial modeling. The natural grassland map, publicly available in GeoTIFF raster format, can be used as a predictor variable for spatial modeling or as a base map for landscape ecology analyses.

* Corresponding author.

E-mail address: laurence.hubert@univ-rennes2.fr (L. Hubert-Moy).

Specification Table

Subject	Nature and Landscape Conservation
Specific subject area	GIS, natural habitat, remote sensing
Type of data	GIS raster and vector
How the data were acquired	The 1770 reference points for natural and artificial grasslands were respectively derived from vegetation maps and the Land Parcel Identification System (LPIS). The 10 m raster of natural grasslands was derived from five annual land cover maps from 2016–2020. All data were produced in a geographical information system using QGIS software [1].
Data format	Analyzed (GIS raster in GeoTIFF format, and vector in GeoJSON format)
Description of data collection	The reference points were collected in 4 steps: <ul style="list-style-type: none"> • Selection of natural and artificial grassland polygons using vegetation maps and LPIS respectively, • Spatial subsampling using 20 × 20 km grids, • Exclusion of polygons < 1 ha, • Visual extraction of one reference point per grassland polygon within an herbaceous area larger than 30 × 30 m (excluding e.g. bare soils and isolated trees) using aerial images. <p>The raster of natural grasslands was created in 2 steps:</p> <ul style="list-style-type: none"> • Calculation of grassland frequency per pixel from 2016–2020, • Classification of natural grasslands (i.e. pixels classified as grasslands every year).
Data source location	Mainland France The raw data used to derive the dataset were provided by regional offices of the environment for the field-based vegetation maps, the French National Geographic Institute (IGN) for the LPIS and aerial images, and the French national land surface data and services hub (Pôle Theia) for the land cover maps.
Data accessibility	Secondary data The dataset of natural grasslands is publicly provided as follows: Repository name: Zenodo Data identification number: 10.5281/zenodo.7895449 Direct URL to data: https://doi.org/10.5281/zenodo.7895449 Primary data The LPIS is released in vector format as follows: Repository name: Geoservices (IGN) Data identification number: IGNF_RPG_2-0 Direct URL to data: https://geoservices.ign.fr/rpg#telechargementrpg2020 Aerial images are released in raster format by French department as follows: Repository name: Geoservices (IGN) Data identification number: IGNF_BDORTHO_r_2-0 Direct URL to data: https://geoservices.ign.fr/bdortho The 2016 land cover map is released in raster format as follows: Repository name: Pôle thématique Surfaces Continentales Theia (CNES) Data identification number: OSO_20160101_RASTER Direct URL to data: https://theia.cnes.fr/atdistrib/rocket/#/collections/OSO/885099e0-aac5-5184-9345-e1b62a55f2ab The 2017 land cover map is released in raster format as follows: Repository name: Pôle thématique Surfaces Continentales Theia (CNES) Data identification number: OSO_20170101_RASTER Direct URL to data: https://theia.cnes.fr/atdistrib/rocket/#/collections/OSO/56364a89-2546-57f3-810b-41deef11e176

(continued on next page)

The 2018 land cover map is released in raster format as follows:

Repository name: Pôle thématique Surfaces Continentales Theia (CNES)

Data identification number: OSO_20180101_RASTER

Direct URL to data: <https://theia.cnes.fr/atdistrib/rocket/#/collections/OSO/c2032fa2-95f3-5dff-866f-b3e2fe5af7d8>

The 2019 land cover map is released in raster format as follows:

Repository name: Pôle thématique Surfaces Continentales Theia (CNES)

Data identification number: OSO_20190101_RASTER

Direct URL to data: <https://theia.cnes.fr/atdistrib/rocket/#/collections/OSO/12e01f87-4e6d-5788-8063-ce59918dae14>

The 2020 land cover map is released in raster format as follows:

Repository name: Pôle thématique Surfaces Continentales Theia (CNES)

Data identification number: OSO_20200101_RASTER

Direct URL to data: <https://theia.cnes.fr/atdistrib/rocket/#/collections/OSO/2327b748-a82c-5933-afb0-087bbfeff4cd>

Value of the Data

- The ground reference points, which accurately geolocate the presence of natural grasslands, were obtained by compiling hundreds of field-based vegetation maps across France.
- The 10 m raster of natural grasslands, based on five years of monitoring, is the first map of such habitats available at high spatial resolution for the whole mainland of France.
- These data can benefit scientists and stakeholders in the fields of ecology, agriculture, or land-use planning.
- The ground reference points can be used as training or test samples for spatial modeling.
- The natural grassland map can be used as a predictor variable for spatial modeling or as a base map in landscape ecology.

1. Objective

Ground reference points for natural grasslands across France were created since no open-source national database was available that compiled hundreds of field-based vegetation maps. A natural grassland map was created because current land cover products at high spatial resolution based on annual satellite time-series do not distinguish natural grasslands from artificial grasslands. The ground reference points were used to validate the accuracy of the 10 m raster of natural grasslands. The map of natural grasslands will be used as a base map to characterize the current fragmentation of grassland habitats in landscapes throughout France. This data paper describes the structure of these data and the method used to generate them for other users and uses.

2. Data Description

The dataset covers for the whole mainland of France (Fig. 1). The ground reference points are provided in standard open-source geospatial data interchange format GeoJSON (`grassland_ground_points.geojson`). These data are projected in the French Lambert-93 projection system (EPSG code 2154) and consist of 882 points for natural grasslands and 888 points for artificial grasslands (total: 1770). Each point is the centroid of a 30×30 m square of herbaceous area, excluding isolated trees, rivers, and impervious areas. The attribute table has two fields: "ID" (the identifier) and "type" (the type of grassland: "natural" or "artificial"). A supplementary layout file (`grassland_ground_points.qlr`) supports formatting of the points in QGIS software.

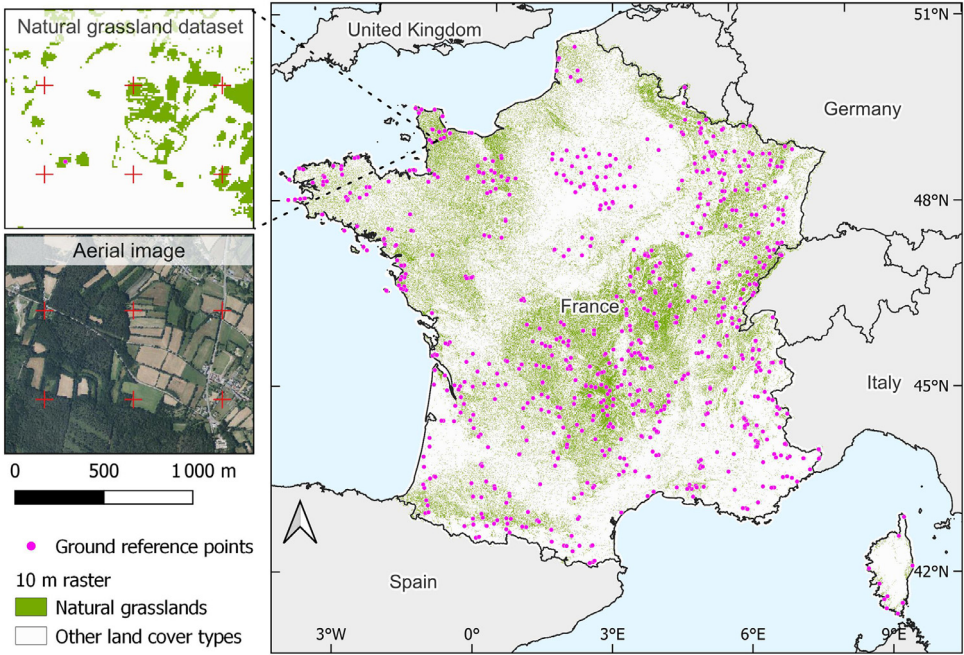


Fig. 1. Ground reference points and pixels of the 10 m raster of the natural grassland dataset across mainland France. For clarity, those of artificial grasslands are not shown.

The map of natural grasslands in 2020 is provided as a file in GeoTIFF format (`natural_grasslands_2020.tif`). It is a categorical raster at 10 m spatial resolution projected in the French Lambert-93 system covering all of mainland France. It includes natural grassland (code 1), but also artificial grassland (code 2) and other land cover types (code 3). A layout file (`natural_grasslands_2020.qml`) supports formatting of the map in QGIS software.

3. Experimental Design, Materials and Methods

Ground reference points for natural and artificial grasslands were extracted from field vegetation maps and the European Union (EU) LPIS, respectively. For natural grasslands, no national open-source database that included all field-based vegetation maps of France was available. Thus, local vegetation maps were requested from each of France’s 13 regional environmental agencies. These vegetation maps cover mainly protected areas [2]. All vegetation maps were then combined in a GIS, and the following polygons, corresponding to natural grasslands, were selected using SQL queries: “Grasslands and lands dominated by forbs, mosses or lichens” (code E and lower nested levels) in the EUNIS classification system [3], “Natural and semi-natural grassland formations” (code 6 and lower nested levels) in the EU Habitat Directive Annex 1 classification system [4], and “Humid grasslands and tall herb communities” or “Mesophile grasslands” (code 37 and 38, respectively, and their lower nested levels) in the CORINE Biotopes classification system [5]. For artificial grasslands, the LPIS for 2020 was used. The LPIS is a polygon vector file that describes the crop type per field based on farmers’ declarations [6]. Polygons classified as “Temporary grassland less than 5 years old” (code 19) were selected.

The polygons of the vegetation maps and LPIS were joined, and only those larger than 1 ha were retained. To balance the spatial distribution of the polygons, sub-sampling was performed

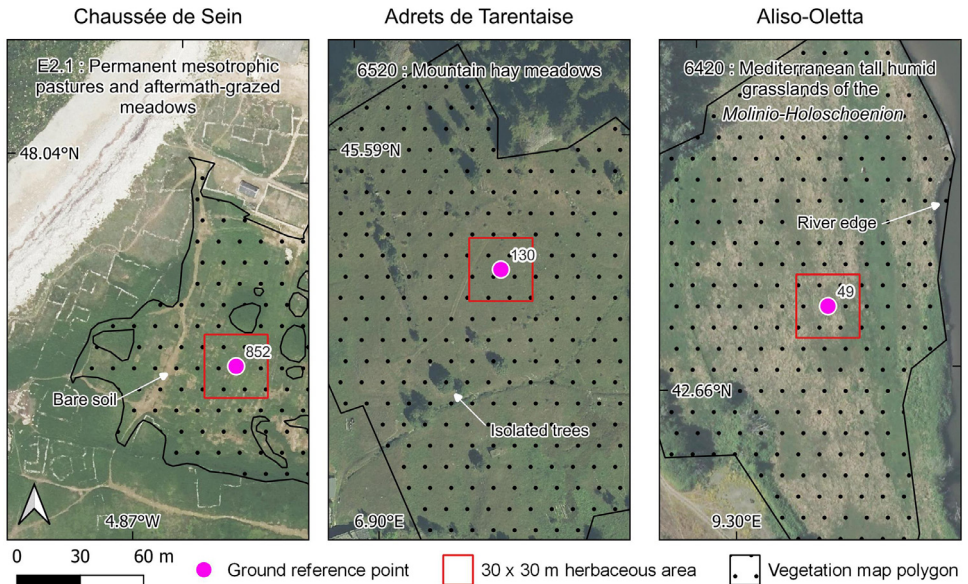


Fig. 2. Examples of selection of natural grassland reference points in three Natura 2000 protected sites: (left) Chaussée de Sein (site code: FR5302007) in the Atlantic biogeographical region, (middle) Adrets de la Tarentaise (site code: FR8201777) in the Alpine biogeographical region, and (right) Aliso-Oletta (site code: FR9400601) in the Mediterranean biogeographical region. The number to the right of each point indicates its identifier.

in 20×20 km grids, with a maximum of 10 polygons per grid for natural grasslands, and a maximum of 1 polygon per grid for artificial grasslands. Additional random subsampling was then performed to obtain a reasonable number (~ 1000 per grassland type) of polygons for visual analysis. Each selected polygon was then visually analyzed in the most recent aerial image, and a reference point was placed at the centroid of each homogeneous herbaceous area that covered at least 30×30 m (Fig. 2). This grid size was set to ensure that the point corresponded to a pure Sentinel-2 pixel (10×10 m) of grassland. Polygons that did not meet the requirement for minimum grassland area (e.g. changes in land cover, narrow patterns) were excluded. Ultimately, 882 and 888 points were collected for natural and artificial grasslands, respectively.

The 10 m raster of natural grasslands was derived from the five annual national land cover maps from 2016–2020 [7–11]. These land cover maps, which had been produced by automatically classifying Sentinel-2 time-series (overall accuracy = 88–91% [12]), were downloaded and stacked into a single raster file. Transition matrix analysis was performed to focus on grasslands (code 211 for 2016–2017, code 13 for 2018–2020) and crops (codes 11–12 for 2016–2017, codes 5–12 for 2018–2020). Pixels classified as grassland every year from 2016–2020 were considered natural grasslands (Fig. 3), while those classified as “crop” at least once were considered artificial grasslands. To support this reasoning, ground reference points were used to assess the accuracy of the classification of natural and artificial grasslands (Table 1). The overall accuracy was high (86%), especially since the few under-detection errors (user’s accuracy) of natural grasslands were related mainly to confusion with non-grassland land cover types.

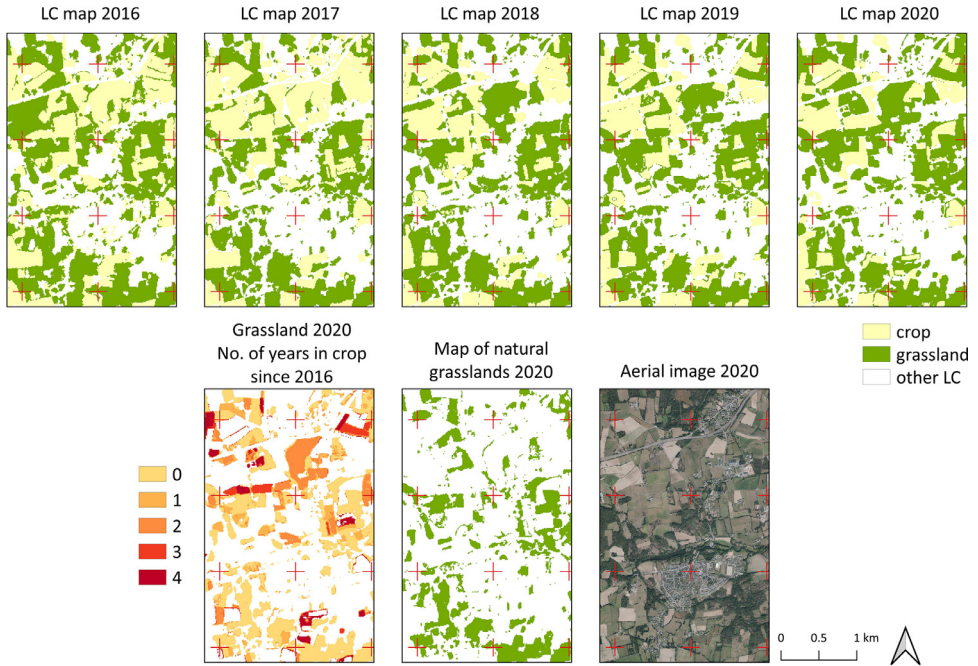


Fig. 3. Illustration of the method used to generate the 10 m raster of natural grasslands in Brittany, France, in the Atlantic biogeographical region (48.03°N, 1.93°W): (top) the land cover (LC) maps from 2016–2020 were combined, (bottom) the number of years the 2020 grasslands had been cropped since 2016 was calculated, and the grasslands that had not been cropped were considered natural.

Table 1

Confusion matrix of the 10 m raster of natural grasslands (prediction) and the ground points (reference). “Others” refers to all non-grassland land cover types. NA = not applicable.

Prediction	Reference			Producer's accuracy
	Others	Artificial grassland	Natural grassland	
Others	0	5	98	0.00
Artificial grassland	0	799	69	0.92
Natural grassland	0	84	715	0.89
User's accuracy	NA	0.90	0.81	
F1-score	NA	0.91	0.85	

Overall accuracy: 0.86.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Natural grasslands across mainland France: a dataset including a 10 m raster and ground reference points (Original data) (zenodo).

CRedit Author Statement

Léa Panhelleux: Formal analysis, Investigation, Software; **Sébastien Rapinel:** Methodology, Supervision, Writing – original draft; **Laurence Hubert-Moy:** Conceptualization, Methodology, Supervision, Project administration, Writing – original draft.

Ethics Statements

Ethics statements do not apply to this dataset since it does not involve the use of human subjects, animal subjects, or social media resources.

Acknowledgments

The authors thank François Chambaud (water agency Rhone-Méditerranée-Corse), Joanie Catrin, Sébastien Abric, Ghislaine Ferrere (French Ministry of Ecology), Guillaume Gayet (OFB), Rachel Vanacker (MNHN), and Anis Guelmami (Tour du Valat) for their productive discussions. The authors are grateful to the 13 regional environment agencies for providing the field vegetation maps. The authors also thank Jean Nabucet (CNRS) for maintaining the IT equipment. This study was supported by the French Ministry of Ecology [grant no. 2103207978].

References

- [1] QGIS Development Team, QGIS Version 2.18. 22. *Geogr. Inf. Syst.* (2018).
- [2] UNEP-WCMC, IUCN, Protected Planet: The World Database on Protected Areas (WDPA) and World Database on Other Effective Area-based Conservation Measures (WD-OECM), (2022). <https://www.protectedplanet.net/> (accessed March 15, 2022).
- [3] C.E. Davies, D. Moss, M.O. Hill, EUNIS habitat classification revised 2004, in: *Report to: European Environment Agency-European Topic Centre on Nature Protection and Biodiversity, 2004*, pp. 127–143.
- [4] Council Directive 92/43/EEC, Conservation of natural habitats and of wild flora and fauna, *Int. J. Eur. Commun.* (1992) 7–49 L206.
- [5] P. Devillers, J. Devillers-Terschuren, J. Ledant, *CORINE Biotopes Manual, Habitats of the European Community, Office for Official Publications of the European Communities, Luxembourg, 1991*.
- [6] European Court of Auditors The Land Parcel Identification System: a useful tool to determine the eligibility of agricultural land – but its management could be further improved, Publications Office of the European Union, LU, 2016 Special report No 25, 2016 <https://data.europa.eu/doi/10.2865/39118> (accessed April 6, 2023).
- [7] V. Thierion, A. Vincent, S. Valero, Theia OSO Land Cover Map 2020, (2022). doi:10.5281/zenodo.6538861.
- [8] Theia OSO Land Cover Map 2019 | Zenodo, (n.d.). <https://zenodo.org/record/6538321#.ZC7QkHvP2Uk> (accessed April 6, 2023).
- [9] J. Inglada, A. Vincent, V. Thierion, Theia OSO Land Cover Map 2106, (2017). doi:10.5281/zenodo.1048161.
- [10] J. Inglada, A. Vincent, V. Thierion, 2017 Metropolitan France Land Cover Map - CESBIO - OSO, (2018). doi:10.5281/zenodo.1993596.
- [11] Theia OSO Land Cover Map 2018 | Zenodo, (n.d.). <https://zenodo.org/record/3613415#.ZC7Qo3vP2Uk> (accessed April 6, 2023).
- [12] J. Inglada, A. Vincent, M. Arias, B. Tardy, D. Morin, I. Rodes, Operational high resolution land cover map production at the country scale using satellite image time series, *Remote Sens.* 9 (2017) 95.