Data in Brief 53 (2024) 110226

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

A spatial dataset about short-term land system dynamics (2005–2015) in the Mediterranean basin

Marta Debolini[a], Elisa Marraccini[b,*]

[a] Fondazione CMCC - Centro Euro-Mediterraneo sui Cambiamenti Climatici, Via Enrico De Nicola 9, Sassari 07100, Italy

[b] Department of Agricultural, Food, Environmental and Animal Sciences (DI4A), University of Udine. Via delle Scienze 206, Udine 33100, Italy

A R T I C L E I N F O

Article history:
Received 13 October 2023
Revised 27 December 2023
Accepted 16 February 2024
Available online 22 February 2024

Dataset link: Mediterranean land system changes between 2005 and 2015 (Original data)
Dataset link: Mediterranean land system classification on 2005 and 2015 (Original data)

Keywords:
Land system dynamics
Intensification
Abandonment
Agricultural and farming systems

A B S T R A C T

Mediterranean land systems are undergoing significant changes in terms of agricultural land use and practices. The location and nature of agricultural areas in the Mediterranean basin are changing following a set of tensions and opportunities taking place both locally and in the entire basin. This dataset presents the main short-term land system dynamics (2005–2015) and their location on the whole Mediterranean basin. The dataset is based on existing land use map and available data about agricultural surface distribution, and it is obtained through a classification process of the land systems on the two analyzed data (2005 and 2015) and a subsequent change detection between the two obtained maps. It covers all the Mediterranean bioclimatic area in pixels of 2 km spatial resolution, harmonizing information from the northern and southern side of the basin. We identified different types of changes, and the most relevant in terms of surfaces are: (1) from mixed agriculture to specialized fruit groves; (2) from agricultural areas to urban and/or periurban areas; (3) from agroforestry to arable systems, and (4) from predominantly bare soils to agricultural areas. This data can be used for further analysis to assess underlying trajectories, and the impact of the observed dynam-

* Corresponding author.
E-mail address: elisa.marraccini@uniud.it (E. Marraccini).
Social media: @MartaDebolini (M. Debolini)

https://doi.org/10.1016/j.dib.2024.110226

ics on biodiversity and the provision of ecosystem services, as well as on the organization of the food system.

Specifications Table

General informaion on the dataset.

Subject area	Agronomy and Crop Science, Global and Planetary Change, Geography, Planning and Development.
More specific subject area	Land system science, Crop distribution, Land system dynamics, Food systems, Agricultural system structure and distribution.
Data format	Analyzed
Type of data	Two shapefiles, one representing the land system classification in 2005 and 2015 and the other representing just the area of changes, with the dynamic description. The shapefiles are structured as a grid of 2 km and for each square of the grid we have a land system type for 2005, a land system type for 2015 and the change between the two data.
Data collection	Online acquisition of land use/land cover map, online acquisition of the SPAM database about crop distribution, cluster analysis of the land systems of the two data, change detection.
Data source location	The dataset represents the whole Mediterranean biogeographical area. The delimitation of the Mediterranean area was based on the combination of the Mediterranean environmental zone proposed by Metzger et al. (2005) for the European part of the Mediterranean, and the limits of Malek and Verburg's (2017) analysis for the southern side of the basin.
Data accessibility	Data provided in the article is accessible to the public at this link: https://doi.org/10.5281/zenodo.8427510 for the shapefile about land system changes and https://doi.org/10.5281/zenodo.8431307 for the land system classification on 2005 and 2015. DOI: https://doi.org/10.5281/zenodo.8431306 and 10.5281/zenodo.8427510

1. Value of the Data

- These data represent the first assessment of land system dynamics on the whole Mediterranean area. This is particularly relevant considering the heterogeneity and the complexity of land systems in the Mediterranean and the difficulties for characterize the agricultural region just in terms of land use.
- The resolution used for presenting Mediterranean land system dynamics (2 km) allows to identify some hotspots of changes.
- This database can be considered as the first assessment needed for evaluating the possible impact of these changes on biodiversity and ecosystem services provision, but also on the food system and food security in the Mediterranean basin.

2. Objective

The spatial patterns of Mediterranean agricultural systems, based on mixed crops and polycultural systems, are difficult to identify and map if only a land cover and land use approach is used [1]. A land system approach, where the geographical unit for modelling and predicting future changes is not directly the land cover or the land use, but a more complex class that combines land use and management through indicators such as population density, irrigation systems, and agricultural inputs is well adapted to tackling Mediterranean agricultural complexity, in order to anticipate and manage the impact of future changes on ecosystem ser-

Table 1
Description of the collected datasets and their sources.

Type of information	Source	Reference	Spatial resolution	Reference time period
Land cover (37 classes, including some management charcateristics such as irrigated and/or rainfed crops)	ESA/CCI base	https://maps.elie.ucl.ac.be/CCI/viewer/	300 m	2005 and 2015
Elevation and slope	Global Digital Elevation Model	https://webmap.ornl.gov/ogc/wcsdown.jsp?dg_id=10003_1	30 Arc-s (∼1 km at the equator)	2015
Soil nutrient availability, rooting conditions and workability	Harmonized world soil database	http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/SoilQuality.html?sb=10	30 Arc-s (∼1 km at the equator)	2008
Temperature and rainfall	Worldclim database	http://worldclim.org/version2	30 Arc-s (∼1 km at the equator)	2010
Population	Landscan	https://landscan.ornl.gov/	30 Arc-s (∼1 km at the equator)	2015
Livestock: amount and type (cattle, goats or sheeps)	Gridded Livestock of the World database	http://www.fao.org/ag/againfo/resources/en/glw/home.html	5 min arc (∼10 km at the equator	2015
Type of crops (33 classes)	SPAM database	http://mapspam.info/"⟩http://mapspam.info/	5 min arc (∼10 km at the equator	2005

vices. However, there is a lack of knowledge on these dynamics for the Mediterranean basin as a whole, as the mapping of land systems dynamics is often fragmented and limited to regional or local case studies. In this context, we explored the current dynamics occurring throughout the Mediterranean basin, aimed at assessing the existing trajectories at a large scale.

3. Data Description

This dataset was originated from the processing of different sources of data, summarized on Table 1. The data were processed through a classification method in order to obtain a land system class on two dates (2005 and 2015). The spatial scale of the two land system maps is 2 km, so each variable was calculated in a 2 km x 2km grid, obtaining a total of 716105 pixels.

As the initial resolution of ESA/CCI land use/land cover database is 300 m (with an overall accuracy of 75% for both years), we chose not to define each 2 km pixel by its dominant land cover class, but to keep the accuracy of the original information by calculating the rate of each land cover class contained in each 2 km square pixel. Each pixel is thus classified, among other variables, according to the proportion (percentage) of each land use/land cover class it contains. The classification thus takes into account land-use variations within the considered spatial units. In order to focus the classification on agricultural areas, we first aggregated some land use classes of the ESA/CCI database, namely the classes related to forests, which was differentiated into broad-leaved and needle-leaved species, wetlands, bare soils (consolidated or unconsolidated) and shrub lands (evergreen or deciduous). We also aggregated some agricultural surfaces of the MAPSPAM database [2], in particular for non-major crops (in terms of their surfaces), such as:

millet, sorghum and other cereals, which were grouped into a single class “other cereals”; bean, chickpea, cowpea, lentil, which were grouped into “pulses”; and soybean, groundnut, sunflower, rapeseed, sesame seed, which were classed as “oil crops”.

4. Materials and Methods

In order to analyze the land system dynamics, we first classified the land system typologies in the Mediterranean basin for 2005 and 2015. We then evaluated the overall changes between the two dates. For the land system classification, we applied a mixed method: we firstly applied an expert-based analysis for the distinction of the non-agricultural classes and then a partitioning of the dataset using an unsupervised classification. For the expert-based analysis, we performed an international survey, submitting an online questionnaire to around 70 scholar and local experts (high level administration, policy makers, land planners) having published on Mediterranean landscape dynamics. The questionnaire consisted of 72 semi-structured questions, including an access to the project data and maps, and organized around five main themes: variables for distinguishing land systems, thresholds for classifying land systems, land systems classes, land system trajectories and underlying drivers. 21 Scholars (30% of the scholars contacted) answered to the questionnaire highlighting a general agreement in the variables for distinguish land systems, and the received answers allowed to identify the thresholds reported on Fig. 1 to distinguish the following non-agricultural classes: forest, wetlands, urban, peri-urban, and bare soils [3]. Then, we performed an unsupervised classification on the remaining pixels, partitioning the whole database in seven subsets based on topographical characteristics (slope and elevation): plain and irrigated plain, hill, hilly plateau, mountain, mountain plateau and high mountain. In each subset we applied the CLARA algorithm – Clustering LARge Applications- created by Kaufman and Rousseeuw [4,5], available in the R package *cluster* [6] and extensively described in [7]. This data mining and clustering algorithm is an extension of the k-medoids methods and deals with large datasets using a sampling approach. CLARA considers a small sample of data with a fixed size and applies the Partitioning Around Medoids algorithm [5] to generate an optimal set of medoids for the sample. The quality of the resulting medoids is measured by the average dissimilarity between every object in the entire data set and the medoid of its cluster, defined as the cost function. CLARA repeats the sampling and clustering processes a pre-specified number of times in order to minimize the sampling bias. The final clustering results correspond to the set of medoids with the minimal cost.

Using CLARA entails specifying the output number of clusters. In order to estimate the optimal number of clusters, we used an unsupervised machine learning approach through the average silhouette method [5,7] contained in the factoextra R package [7]. This method directly determines the optimal number of clusters k by measuring the quality of the clustering, i.e. how well each object lies within its cluster. It computes the average silhouette of observations for different values of k. The optimal number of clusters k is thus the one that maximizes the average silhouette over a range of possible values for k [5].

For the clustering procedure, we used all the data acquired as variables, without applying any additional threshold, i.e population, crop surface, land use, livestock, soil quality, temperature and rainfall.

Fig. 1 shows the hierarchical scheme applied for the land system classification for the two dates.

From the land system classification, we obtained 20 classes. Fig. 2 shows the distribution of the 20 classes on the Mediterranean for the year 2015. For validating the obtained classification, the resulting maps were showed and discussed with local partners (at least three expert scholars) of seven Mediterranean countries (Portugal, Spain, Italy, France, Tunisia, Malta and Algeria). Moreover, the expert survey previously described was also used to validate the classes, obtaining a substantial agreement for the typology (data not shown).

The most represented classes are mountain plateau with extensive agriculture (around 362,000 km^2, 15% of the total area) and moderately intensive plain characterized by cereals and

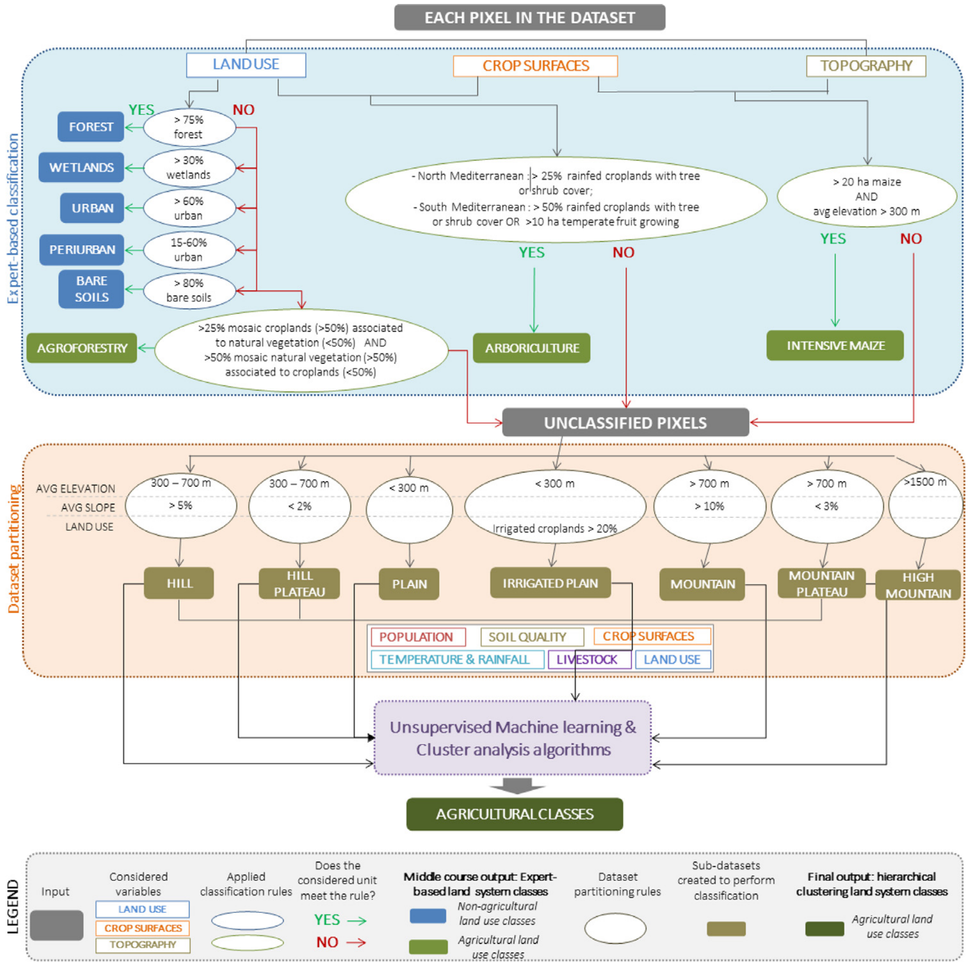

Fig. 1. Overall methodology for land system classification.

arable lands (around 250,000 km^2, 10% of the total area), showing that mixed systems are still prevalent in the Mediterranean region. Intensive agriculture is also generally diffused both on the northern and southern side of the Mediterranean with around 190,000 km^2. In general, intensive systems occupy around 23% of the total surfaces, whereas extensive systems are around on 50% of the areas.

Then, we assessed the changes between 2005 and 2015, representing around 57,000 km^2, 2.5% of the total area. The confusion matrix allows to identify most frequent dynamics:

(1) From agricultural areas to urban and/or periurban areas (~43% of the changing area). The highly productive plains with a predominance of cereal crops or hill plateau areas with arable crops or fruit groves are the most affected by this kind of change (36% and 6%, respectively) together with arboriculture areas (10%). This means that Mediterranean areas are losing their most productive agricultural lands to urbanization. These areas are widespread around the Mediterranean basin, mainly in the north-western and south-eastern parts.

(2) From mixed systems to specialized fruit groves (~28% of the changing area). This corresponds to a change from different types of land systems including arable crops, vegetables

Fig. 2. Land system class distribution in 2015.

and livestock, to specialized fruit groves. These dynamics take place in different types of topography, from plains to mountains. We observed these dynamics in the Nile plain, in some coastal areas of Turkey, but also in Albania, southern Italy (Apulia), Valencia (Spain) and in some coastal areas of the Maghreb.

(3) From extensive mixed systems and/or agroforestry to more intensive systems (~4% of the changing area). This highlights the simplification of agroforestry systems or the extensive arable/pastures through the loss of trees and the increase in arable lands, until the complete disappearance of agroforestry and the transition to arable crops and vegetables. The areas most affected by these changes are in the internal parts of Portugal, some coastal areas of Algeria, and the internal hilly parts of the Balkans and Turkey.

(4) From predominantly bare soils to agricultural areas (~11% of the changing area). There has been a change from predominantly bare soils to extensive agricultural areas. It has especially affected the southern Mediterranean, particularly Morocco and Algeria, near the borders between agricultural and desert areas, or in the southern Atlas Mountains.

Fig. 3. Land system change distribution in the Mediterranean basin from 2005 to 2015.

(5) Agricultural abandonment and re-naturalization (~7% of the changing area). This includes all the transitions from an agricultural system class to the forest and natural vegetation type. The areas most affected by this change are the internal hilly and mountain lands of the northern side of the Mediterranean.
(6) From forest and/or natural vegetation to agricultural areas (~5% of the changing area). This represents all the transition from the forest class to an agricultural system. This type of change is mainly present on the mountain areas of the northern side of the Mediterranean, and in particular in some valleys in Slovenia and of the Italian Alps (Fig. 3).

Ethical Statement

Hereby, we consciously assure that for the manuscript "A spatial dataset about short-term land system dynamics (2005-2015) in the Mediterranean basin" the following is fulfilled:

(1) This material is the authors' own original work, which has not been previously published elsewhere.
(2) The paper is not currently being considered for publication elsewhere.
(3) The paper reflects the authors' own research and analysis in a truthful and complete manner.
(4) The paper properly credits the meaningful contributions of co-authors and co-researchers.
(5) The results are appropriately placed in the context of prior and existing research.
(6) All sources used are properly disclosed (correct citation). Literally copying of text must be indicated as such by using quotation marks and giving proper reference.
(7) All authors have been personally and actively involved in substantial work leading to the paper, and will take public responsibility for its content.
(8) This work does not involve studies with animals and humans

Data Availability

Mediterranean land system changes between 2005 and 2015 (Original data) (Zenodo).
Mediterranean land system classification on 2005 and 2015 (Original data) (Zenodo).

CRediT Author Statement

Marta Debolini: Conceptualization, Methodology, Formal analysis, Writing – original draft; **Elisa Marraccini:** Conceptualization, Validation, Writing – original draft.

Acknowledgments

The data presented were produced on the frame of the DIVERCROP project, founded through the ARIMNet2 2016 Call by the following funding agencies: ANR (France), IRESA (Tunisia), INIA (Spain), FCT (Portugal), ATRSNV (Algeria), MIPAAF (Italy) and MCST (Malta). ARIMNet2 (ERA-NET) has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 618127. Special thanks to all of the local partners of the DIVERCROP project for collaborating on data collection, discussing the method and validating the results.

We acknowledge the contribution of Johanna Fusco for the data management in the initial version of the database and of Mina Sadeghzadegh for her contribution to the expert survey.

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.

References

[1] Ž. Malek, P. Verburg, Mediterranean land systems: representing diversity and intensity of complex land systems in a dynamic region, Landsc. Urban Plann. 165 (2017) 102–116.
[2] You, L., U. Wood-Sichra, S. Fritz, Z. Guo, L. See, and J. Koo. (2017). Spatial production allocation model (SPAM) 2005 v3.2.October 17, 2018. Available from http://mapspam.info.
[3] E. Marraccini, M. Sadeghzadegh, M. Debolini, Lessons learnt from an expert-based validation of land system dynamics (2005–2015), in: the Mediterranean basin. IALE22 European Landscape Ecology Congress, Warsaw Poland, 2022, pp. 11–15. July 2022.
[4] Kaufman, L., and Rousseeuw, P. J. (1986), Clustering Large Data Sets (with discussion), in Pattern Recognition in Practice II, edited by E. S. Gelsema and L. N. Kanal, Elsevier/North-Holland, Amsterdam, pp. 425-437.
[5] L. Kaufman, P. Rousseeuw, Finding Groups in Data: An Introduction to Cluster Analysis, Wiley, 1990 http://www.sthda.com/english/articles/27-partitioning-clustering-essentials/89-clara-clustering-large-applications/.
[6] Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K. (2018). Cluster: cluster analysis basics and extensions. R package version 2.0.7-1.
[7] Kassambara A. (2017). Practical guide to cluster analysis in R. Unsupervised machine learning. In Multivariate analysis I, edited by sthda.com, 186pp.