Final LanDS toolbox

REACT4MED Project (Grant Agreement No. 2122) is funded by PRIMA - Partnership for Research and Innovation in the Mediterranean Area, a Programme supported by Horizon 2020.

This page is intentionally left blank

Technical References		
Project Acronym	REACT4MED	
Project title	Inclusive Outscaling of Agro-ecosystem REstoration ACTions for the MEDiterranean	
Project coordinator	HMU	
Project Duration	36 months	
Deliverable No.	D4.3	
Deliverable title	Final LanDS toolbox	
Lead partner	SoftWater s.r.l.	
Contributing partners	TUC, UH, UOS, CIHEAM Bari, HMU	
Author(s)	Davide Cantaluppi, Simone Corti, Elena Matta, Marco Micotti, Matteo Sangiorgio, Enrico Weber (SOFTW)	
Type	Document	
Format	MS-Word	
Language	EN-GB	
Creation date	6 December 2024	
Version number	V01.2	
Version date	10 January 2025	
Last modified by	Elena Matta	
Due date	31 December 2024	
Actual delivery date	10 January 2025	
Rights	Copyright © 2022, REACT4MED Consortium	
Dissemination level		CO (confidential, only for members of the consortium)
	x	PU (public)
		PP (restricted to other programme participants)
		RE (restricted to a group specified by the consortium)
	When restricted, access granted to:	
Nature	x	R (report)
		P (prototype)
		D (demonstrator)
		O (other)

Revision history			
Version	Date	Modified by	Comments
0.0	6 December 2024	Elena Matta	First table of content's draft of the report
1.0	30 December 2024	Davide Cantaluppi, Simone Corti, Elena Matta, Marco Micotti	First draft of the report
1.1	8 January 2025	Elena Matta, Marco Micotti, Matteo Sangiorgio	Revised draft of the report
1.2	10 January 2025	Elena Matta, Marco Micotti, Enrico Weber	Refined and submitted draft of the report

Table of Contents

List of acronyms

CDD: Maximum Length of Dry Spell
CID: Climate Impact Drivers
CIHEAM Bari: Istituto Agronomico Mediterraneo di Bari
CLI: Command Line Interface
CY: Cyprus
CyI: The Cyprus Institute
CMS: Content Management System
D: Deliverable
DB: Database
DM: Decision Maker
DoA: Description of Action
EG: Egypt
ES: Spain
ERLL: Ecosystem Restoration Living Lab
ETCCDI: Expert Team on Climate Change Detection and Indices
GA: Grant Agreement
GPU: General Public Licence
GR: Greece
GWLs: Global Warming Levels
HMU: Hellenic Mediterranean University
IDF: Intensity Duration Frequency
IL: Israel
INRA: Institut National de la Recherche Agronomique
IPCC: Intergovernmental Panel on Climate Change
IT: Italy
ISIMIP: Integrated Science Impact Model Intercomparison Project
IVS: Input Variable Selection
JSON: JavaScript Object Notation
LanDS: Land degradation Decision-Support
M: Month
ML: Machine Learning
MO: Morocco
MR: Matching Rate
MS: Milestone
NetCDF: Network Common Data Form
NDII: Normalized Difference Infrared Index
NDWI: Normalized Difference Water Index
NSDI: Normalized Difference Soil Index
NGO: Non-Governmental Organisation
PA: Pilot Area
PAL: Pilot Area Leader
PCA: Principal Component Analysis
PDS: Participatory Development Solutions
PPs: Project Partners
PrAvg: Daily average precipitation
PRIMA: Partnership for Research and Innovation in the Mediterranean Area

PSRI: Plant Senescence Reflectance Index
RCPs: Representative Concentration Pathways
SAVI: Soil Adjusted Vegetation Index
SH: Stakeholder
SIPI: Structure Insensitive Pigment Index
SOFTW: SoftWater s.r.l.
SSPs: Shared Socioeconomic Pathways
STAC: SpatioTemporal Asset Catalogs
TR: Turkey
TUC: Technical University of Crete
UH: University of Haifa
UOS: Osnabrück University
UTAEM: Turkish International Agricultural Research and Training Center
UV: Universidad de Valencia
WFS: Web Feature Service
WMS: Web Map Service
WP: Work Package
WS: Workshop
WSDI: Warm Spell Duration Index

Executive Summary

Deliverable D4.3 describes the final version of the LanDS toolbox,available at the following URL: http://lands.soft-water.it, composed of four tools: (1) a geo-referenced data repository serving as a knowledge base by collecting site-specific data and resources from the ERLLs (WP3) as well as broader scale information from global or regional public repositories and satellite-based indices (WP2); (2) a data viewer, containing a set of visual analytics tools linked with this repository allowing effective data access and sharing among project partners and stakeholders, and interactive visualizations, supporting the monitoring of restoration actions (WP5) and dissemination of project outcomes (WP7); (3) an indicators library implemented as a modular and generalised code library, applicable to different geographical contexts based on collected data and indicators identified in collaboration with WP2 and WP3; (4) a machine-learning based procedure to identify potentially suitable areas in the Mediterranean for up- and outscaling of restoration measures. A fifth tool (5), namely the LanDS Dashboard, due in M34 and under development, will provide a user-friendly and interactive interface, supporting policy makers and other stakeholders in the land degradation assessment and evaluation of impacts of restoration measures. Combining existing knowledge from global and regional renowned repositories and expertise arisen from the living labs in the pilot areas, the LanDS is aimed at supporting better-informed land restoration actions and more sustainable resources management, enabling – at the final stage of the project – to elaborate policy recommendations and to identify investment opportunities for public and private actors, based on the criteria of maximum cost-effectiveness and impact (in collaboration with WP5 and WP6).

1 Introduction

REACT4MED aims to extend the potential application of the land restoration actions promoted in the Ecosystem Restoration Living Labs (ERLLs) running in the different pilot areas (PAs) to the Mediterranean scale. Within the project, Work Package 4 (WP4, Science-based decision support toolbox) efforts are focused on developing a scientific Land degradation Decision-Support Toolbox (LanDS) and applying it at different spatial and temporal scales. The LanDS development is driven by a co-creation approach, involving stakeholders (SHs) and decision-makers (DMs) from ERLLs in the design and evaluation of the toolbox, to ensure portability and maximise the effectiveness of the toolbox. This involves a tight collaboration with WP3 and PA leaders, who are our first SHs and information sources in the different study areas, but also with WP2, which reviews and provides significant literature and broad-level biophysical and socio-political indicators, including satellite-based land productivity and climatic trajectories, on which the LanDS also builds on. An overview of the REACT4MED WPs interaction and planning during the project's lifetime is shown in Figure 1 (from the Description of Action - DoA, i.e., Annex I of the GA).

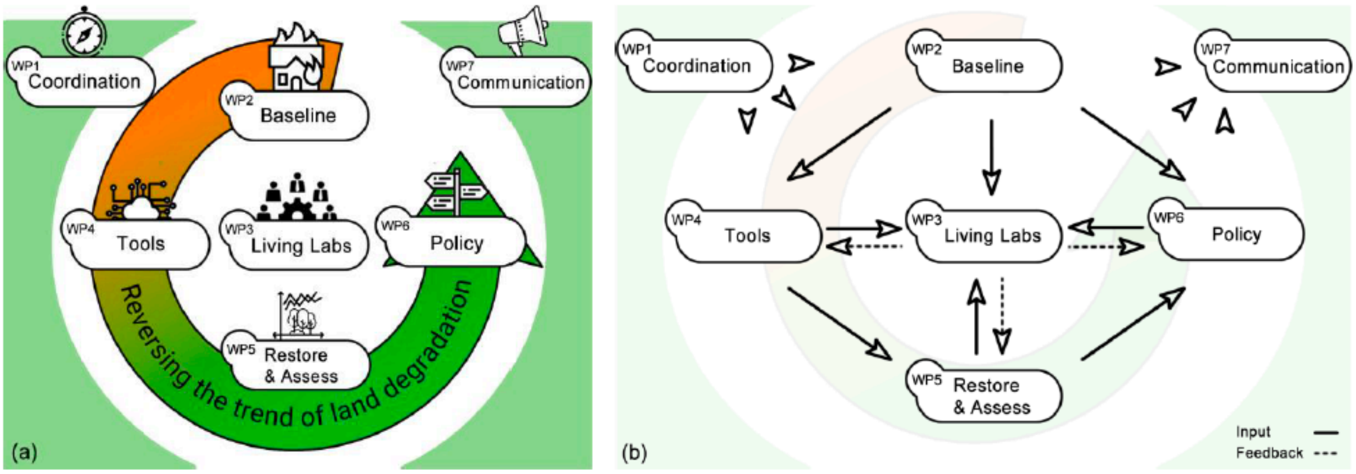


Figure 1 REACT4MED Work Package (WP) structure (a) and information flow within its conceptual framework (b).

In this framework, the multi-operational LanDS toolbox is aimed at supporting a harmonized land degradation assessment and impact evaluation of REACT4MED restoration measures, applicable to different geographical and social contexts. This deliverable consists in the accompanying document of the final release of the LanDS toolbox, including the planned tools, populated with project's data, and documented by an online user-guide with examples and tutorials[1]. The toolbox is deployed and accessible at LanDS website[2] while software developed is released in an online public repository[3].

The scheme of the LanDS toolbox, already presented in detail in D4.1 and D4.2, is shown in Figure 2. The five tools included in the LanDS and correspondent tasks which are and will be performed are the following: (1) a geo-referenced data repository serving as a knowledge base by collecting site-specific data and resources from the ERLLs (WP3) as well as broader scale information from global or regional public repositories and satellite-based indices (WP2); (2) a data viewer, containing a set of visual analytics tools linked with this repository allowing effective data access and sharing among project partners and stakeholders, and interactive visualizations, supporting the monitoring of restoration actions (WP5) and dissemination of project outcomes (WP7); (3) an indicators library implemented as a modular and generalised code library applicable to different geographical contexts based on collected data and indicators identified in collaboration with WP2 and WP3; (4) a machine-learning based procedure (alias ML tool) to identify potentially suitable areas in the Mediterranean for up- and outscaling of restoration measures. The fifth tool

---

[1] https://lands.soft-water.it/documentation
[2] http://lands.soft-water.it
[3] https://gitlab.com/lands-r4m

(5) is the LanDS web interactive dashboard, which is still in development and will provide a harmonised land degradation assessment and evaluation of impacts of restoration measures, as well as a user-friendly interface (the correspondent deliverable D4.4 is due in M34). In a final phase, plausible combinations of future climate and land-use change, and socio-economic scenarios will facilitate the detection of potential degradation trends in medium and long-term time horizons, providing useful insights to orient policy recommendations and support decision-making processes (WP5 and WP6).

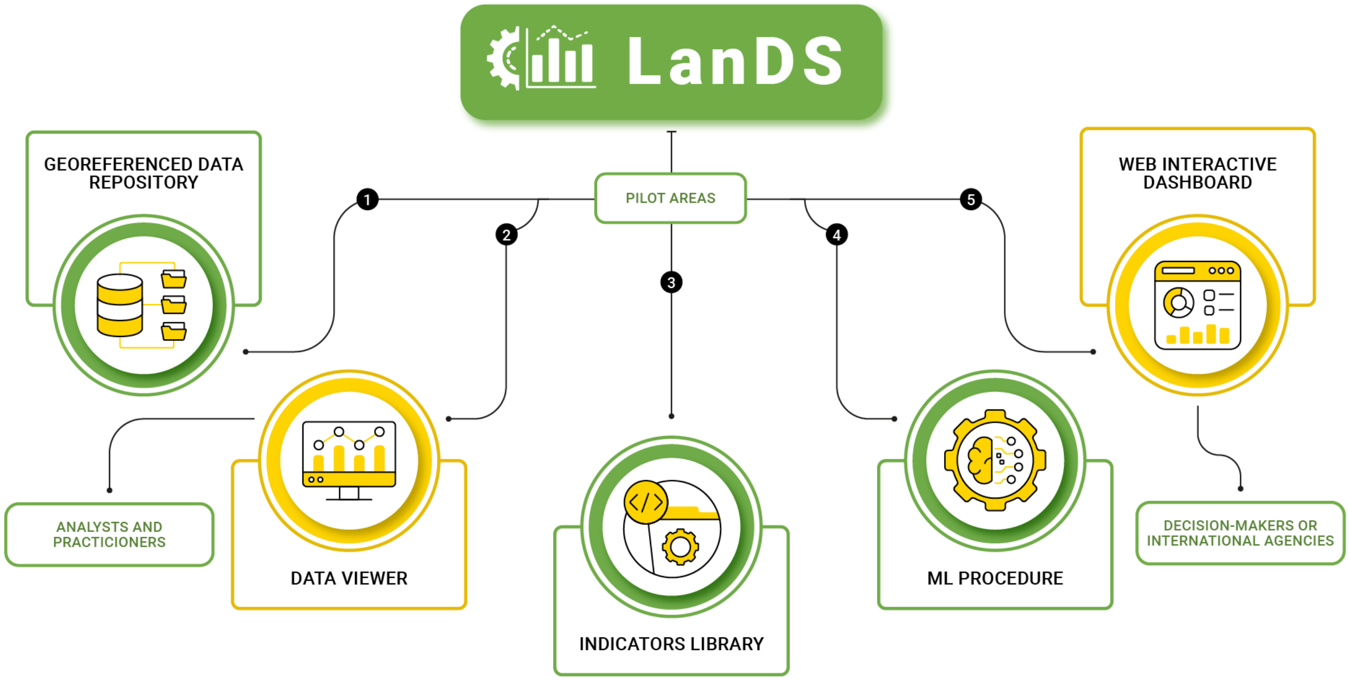


Figure 2 Scheme of the LanDS toolbox, composed of five main tools, to be developed by the end of the project (April 2025).

The document is structured as follows: after this introduction, Section 2 describes the LanDS toolbox, outlining the main characteristics and functionalities of its four tools: the repository, organised in map, collections, and data browser; the data viewer; the indicators library; and the ML tool. Section 3 briefly presents the online documentation (user-guide), while Section 4 summarized the open-source stack of technologies implemented. We conclude with some final remarks and a brief overview of the next steps (Section 5). The References and Annexes complete the document.

2 The LanDS toolbox

The LanDS toolbox is composed of four different tools performing different tasks, as shortly presented in the introduction and schematized in Figure 2: (1) a geo-referenced data repository; (2) a data viewer, (3) an indicators library implemented as a modular and generalised code library, and (4) a ML tool. The toolbox is available at the LanDS URL[4], which is also available from REACT4MED website [5], while the source code has been released with an open-source licence in a public code repository[6].
In this Section, we will present the main features and snapshots of the four tools, that can be reached from the top bar menu of the LanDS web interface, while, on the right-hand side, there is the link to the documentation (indicated by an "i") and login (useful to upload and/or edit data items).

2.1 Repository

The first released tool of the LanDS consists in a geo-database and a structured repository of spatial data, aimed to build a knowledge base by collecting site-specific data and resources from the ERLLs as well as broader scale information from global or regional public repositories and web services. It is accessible from the top bar menu of the LanDS web interface, and it is organized in Map, Collections, and Data Browser. The LanDS data repository is compliant with STAC[7] specification, a common language to describe geospatial information, so it can more easily be worked with, indexed, and discovered. Consequently, each dataset stored in the repository is described as a *STAC Item,* the core atomic unit of STAC, representing a single spatiotemporal asset as a GeoJSON feature plus datetime and links, while set of datasets are represented as *STAC Collection*.

2.1.1 Map

Figure 3 shows the LanDS landing page, presenting the map of the Mediterranean Basin with nine location pins to highlight the nine available collections (one for each REACT4MED PA plus a general one for the Mediterranean Area) and, on the right-hand side, a column with the collections represented by their reference image, their title/location, and associated issues (e.g., soil erosion), which can be easily scrolled down and selected as needed. Thus, a user has the option to choose the collection *i* from the pins on the map or for the reference images.

---

[4] http://lands.soft-water.it/
[5] https://react4med.eu/
[6] https://gitlab.com/lands-r4m
[7] https://stacspec.org/

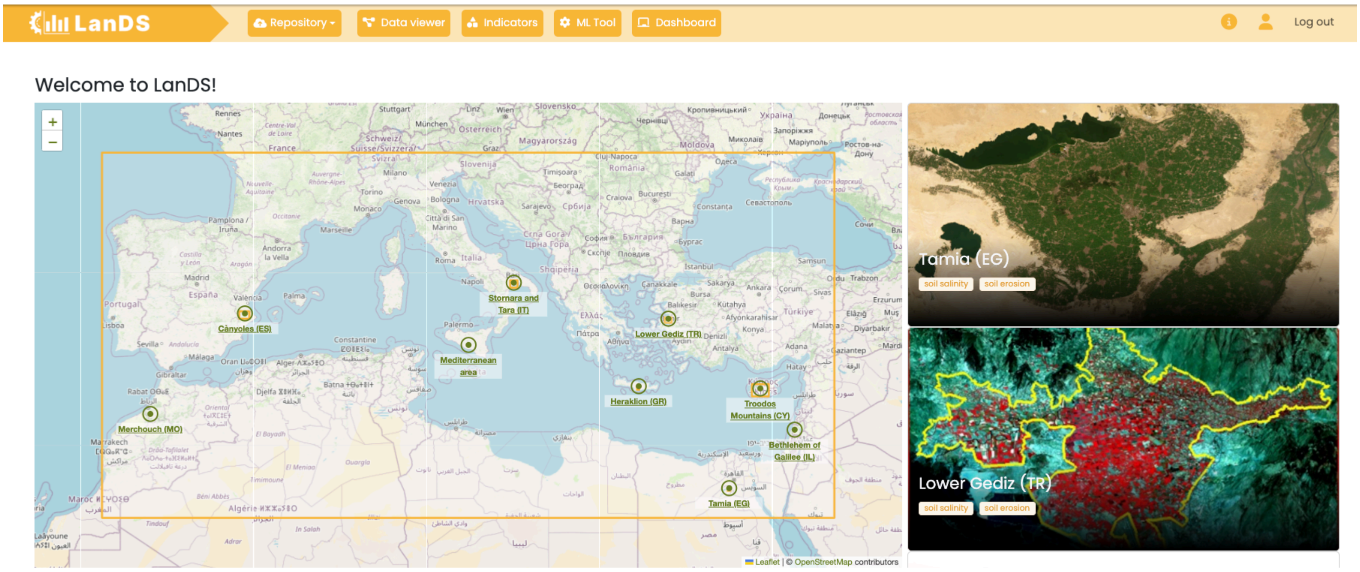


Figure 3 LanDS landing page.

2.1.2 Collections

Figure 4 List of LanDS Collections: in this screenshot, four out of nine available collections are shownshows the LanDS collections, displayed with their reference image, their title/location, and their tags (i.e., issues and/or restoration actions). The collections are then organised in the data browser (Section 2.1.3), where data can be stored as project's resource and/or computed indicator as list of datasets (items) and files and presented with a title (also used for filtering or sorting).

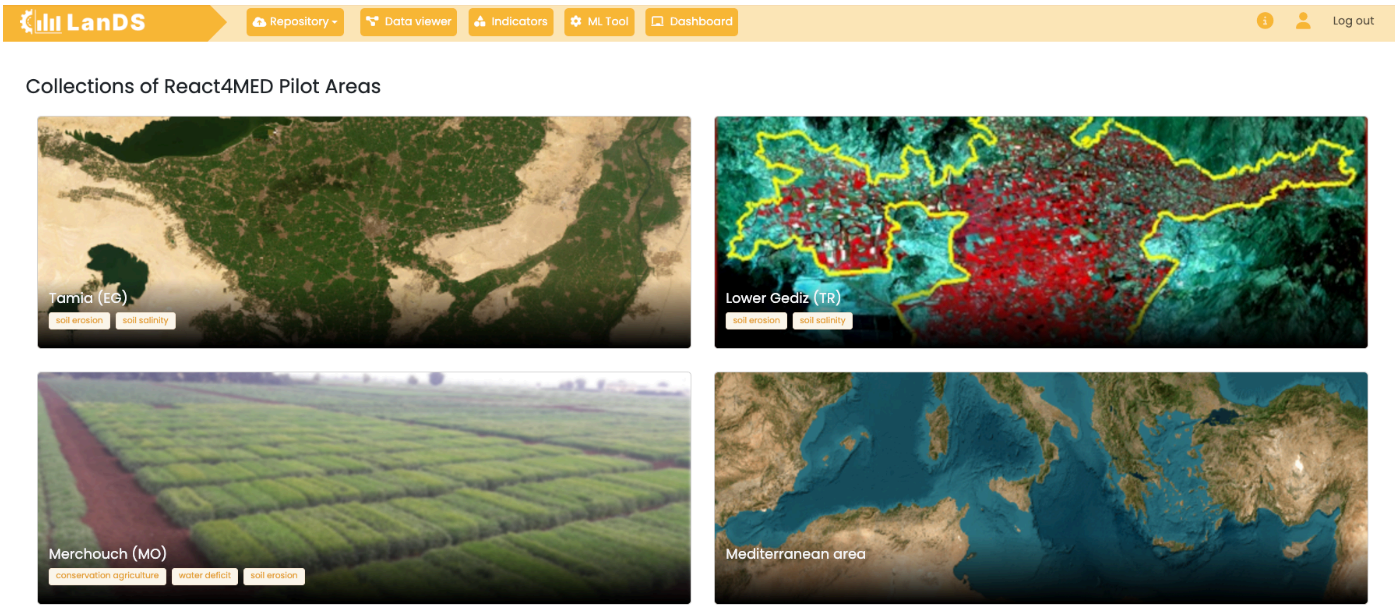


Figure 4 List of LanDS Collections: in this screenshot, four out of nine available collections are shown.

By clicking on the title of a specific collection (from Map, Collections, or Data Browser), the collection page is displayed. Figure 5 and Figure 6 present an example of the collection page for PA1 in Cyprus (Troodos Mountains) for two categories of information: Info & Media and Dataset, respectively. They include:

- Title, description, keywords, and any available documentation organized under Media & Documentation (Figure 5, left-hand side);

- Dataset (Figure 5, left-hand side), containing all information related to this collection, which can be classified as different types: Resources or LanDS indicators (see subsection 2.3) and are displayed with their title, description, temporal extent, and files;
- Bounding box/spatial extent framing the dataset location (Figure 5 and Figure 6, right-hand side).

If the user is logged in, as in the example reported Figure 6, next to the collection title there is the possibility to click on *Edit collection* by adding or removing information included in the Info & Media and Dataset sections.

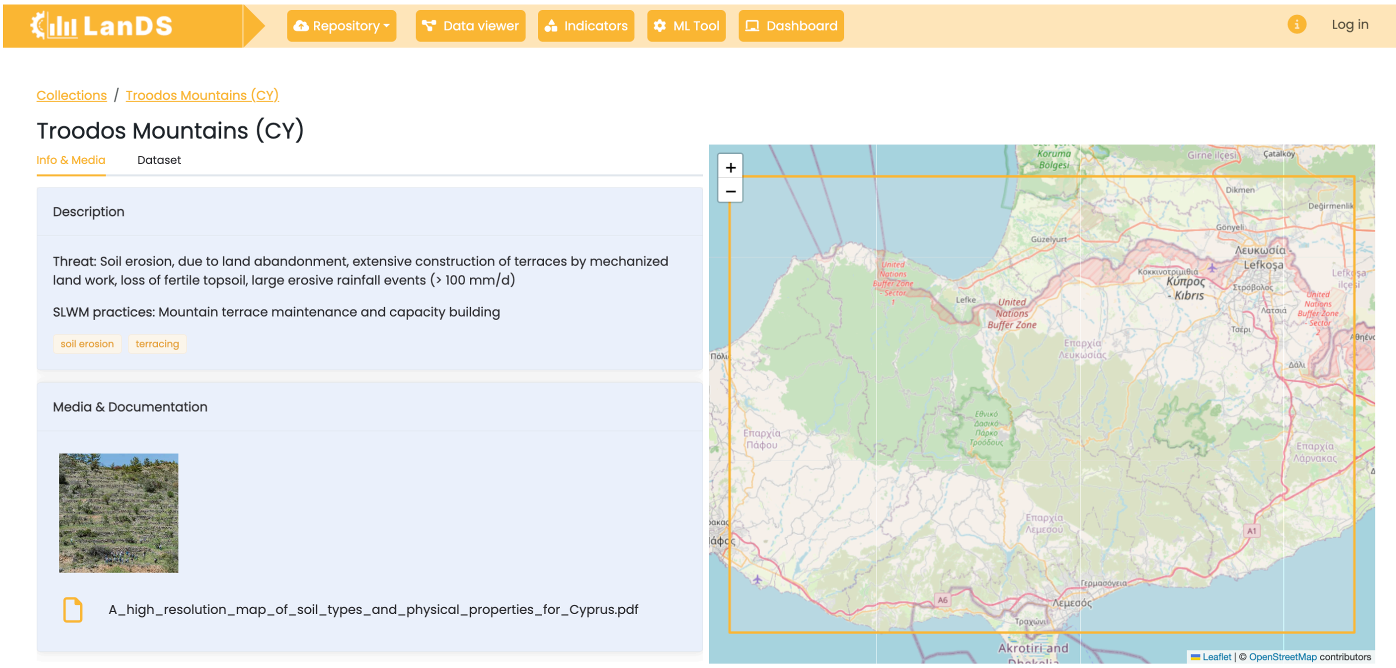


Figure 5 Example of the Cyprus - Troodos Mountains (PA1) collection page, showing the Info & Media section (left-hand side) and the bounding box/spatial extent (right-hand side).

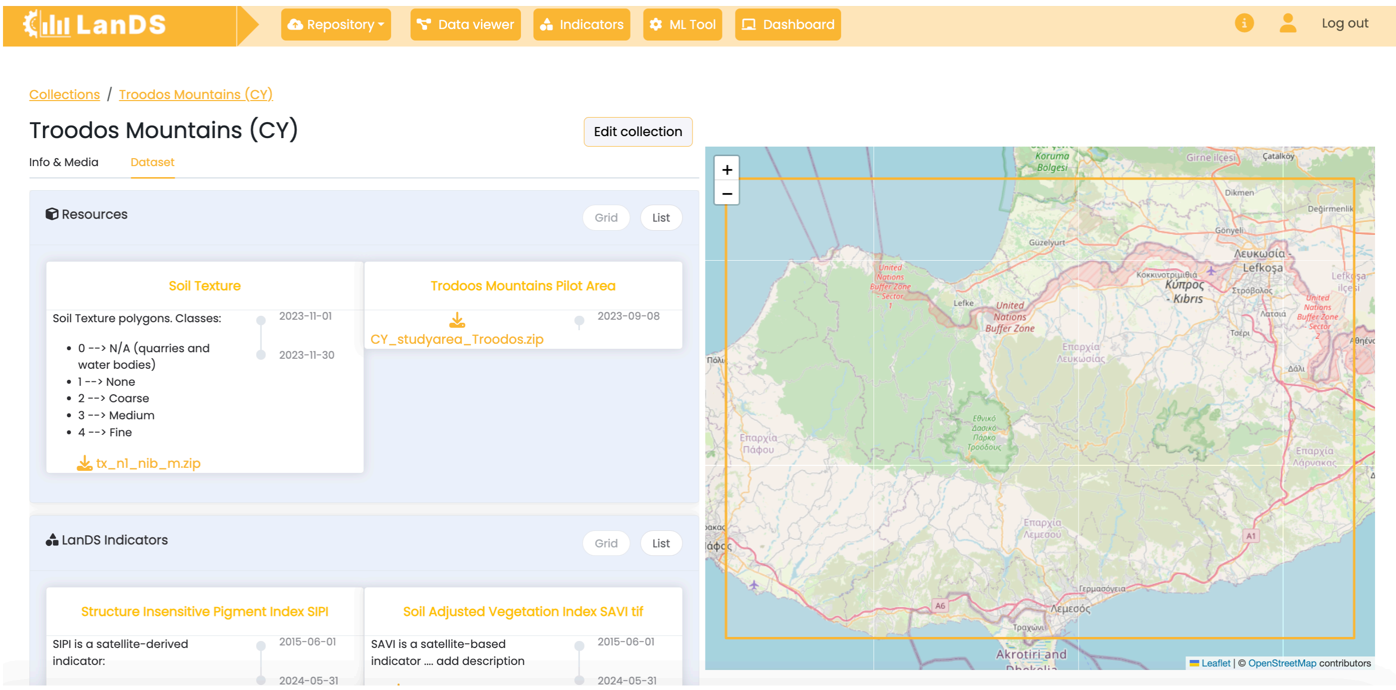


Figure 6 Example of the Cyprus - Troodos Mountains (PA1) collection page, showing the dataset section (left-hand side) and the bounding box/spatial extent (right-hand side).

2.1.3 Data browser

The data browser page (Figure 7) provides a quick access to the datasets in LanDS repository. All data stored in the repository is linked either to a specific case study or to the whole Mediterranean Area and, thus, organized in nine available Collections: one for each REACT4MED PA plus a general one for the Mediterranean Area.

The data browser consists in a dynamic table where datasets are grouped by belonging collections and linked to their descriptive page (by clicking on the dataset name) and to their direct download (by clicking on the icon on the right-hand side of the page. Other available functionalities are the following ones:

- Show: it displays the selected number of available entries;
- Search: it provides a quick filter on displayed rows based on text included in each field;
- Advanced Search: it performs a more refined spatio-temporal query across all the collections.

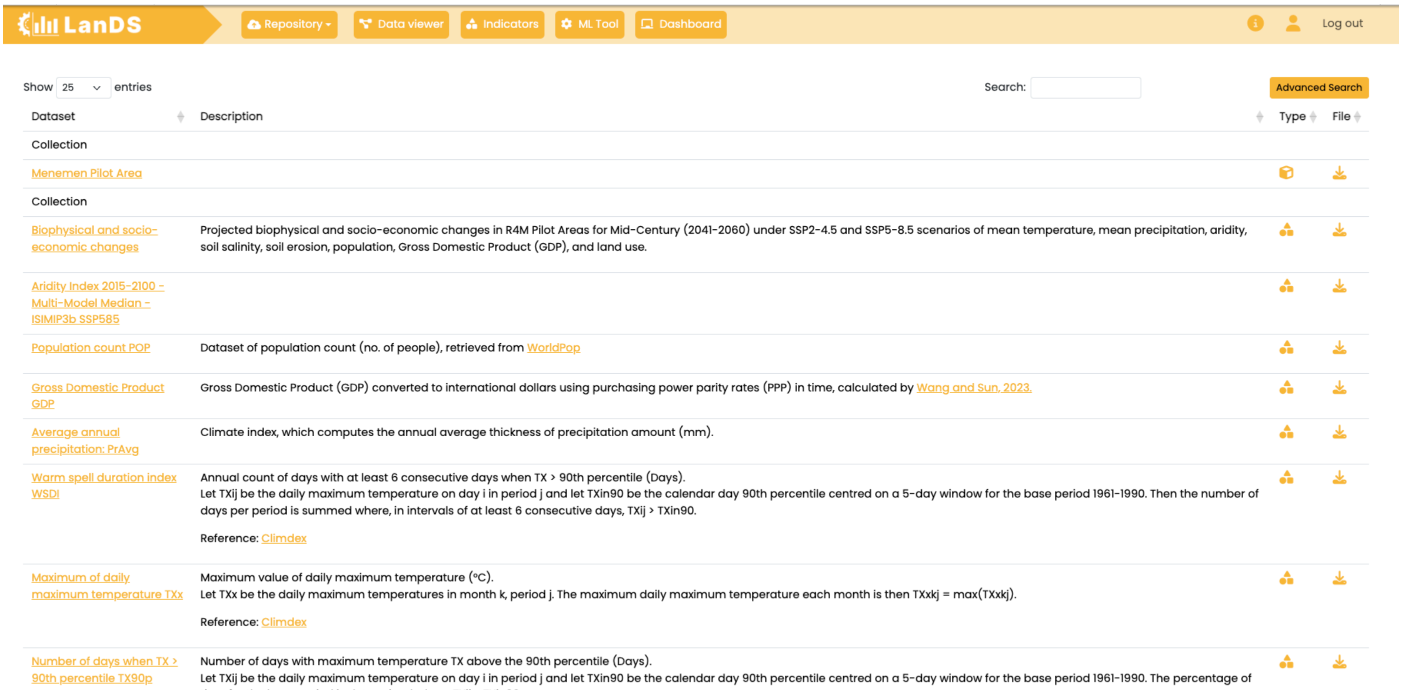


Figure 7 Screenshot of the data browser page, showing a part of the stored datasets, along with its functionalities (show, search, advanced search) and characteristics (resource or indicator), with the possibility of downloading them (on the right-hand side of the page).

2.1.3.1 Dataset page

By clicking on the title of a specific dataset on the data browser, the related page is displayed. Figure 8 shows an example for a soil salinity dataset for the Mediterranean Area, with the same general functionalities already described above and the following specific contents and functionalities:

- Title of the dataset and description (if available), temporal extent and type of data (left-hand side of Figure 8);
- Data preview (automatically shown): Data Cube section presents the metadata included in the NetCDF file (.nc), where all available information is added such as variables and dimensions (right-hand side of Figure 8);
- Spatial extent (it should be clicked on in order to be shown): bounding box of the dataset;
- Direct download link (*Download* button) and possibility to be visualized in the data viewer (*Open in data viewer* button, see subsection 2.2);
- Possibility to edit the dataset (item), if the user is logged in.

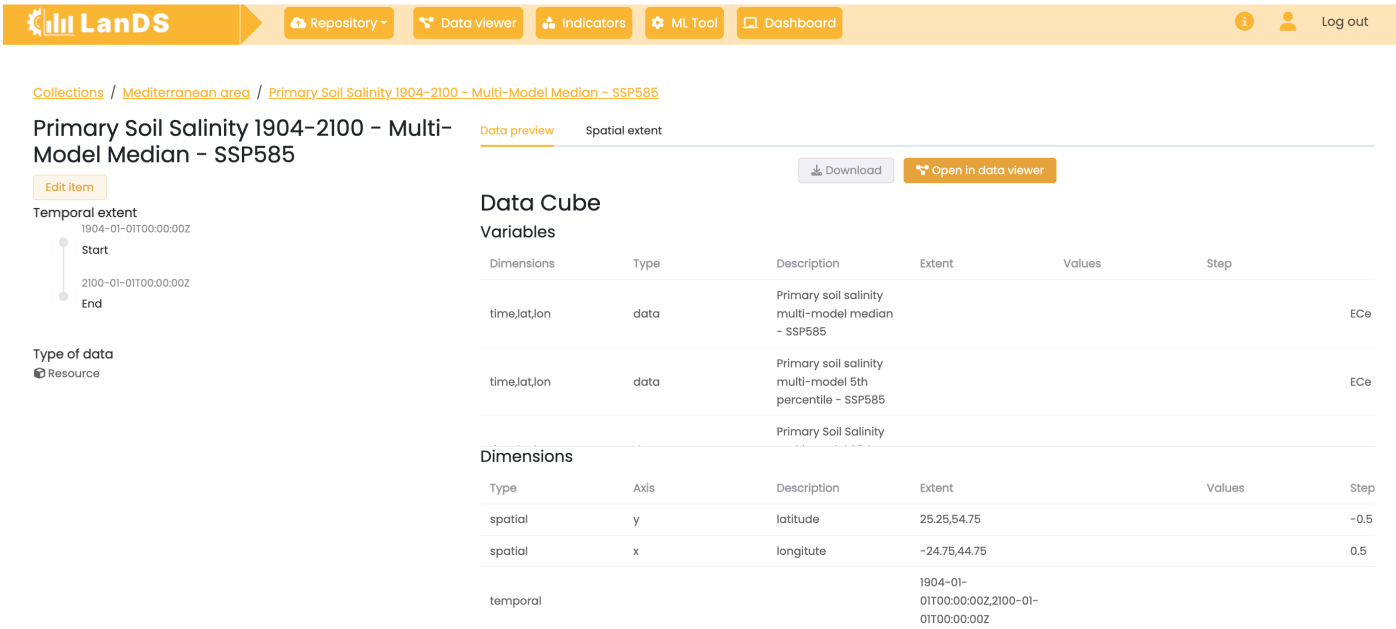


Figure 8 Example of a dataset (item) detail page: in this case, showing Soil Salinity for scenario SSP585 in the Mediterranean Area.

Figure 9 shows an example of a data upload form, available when the user is logged in and selects the correspondent collection, where each dataset (item) can be uploaded following three steps:

- Step 1: the dataset (item) can be described with a title (item name), a rich text description, one or more files uploaded through the *upload file* button;
- Step 2: to define the temporal extent;
- Step 3: to add a certain number of metadata referred to the DataCite standard schema[8], with different object properties that can be inserted (e.g., publisher, identifier, creator), although, it is not mandatory to be able to save the item (there is a *Save item* button at the end of Step 3 section).

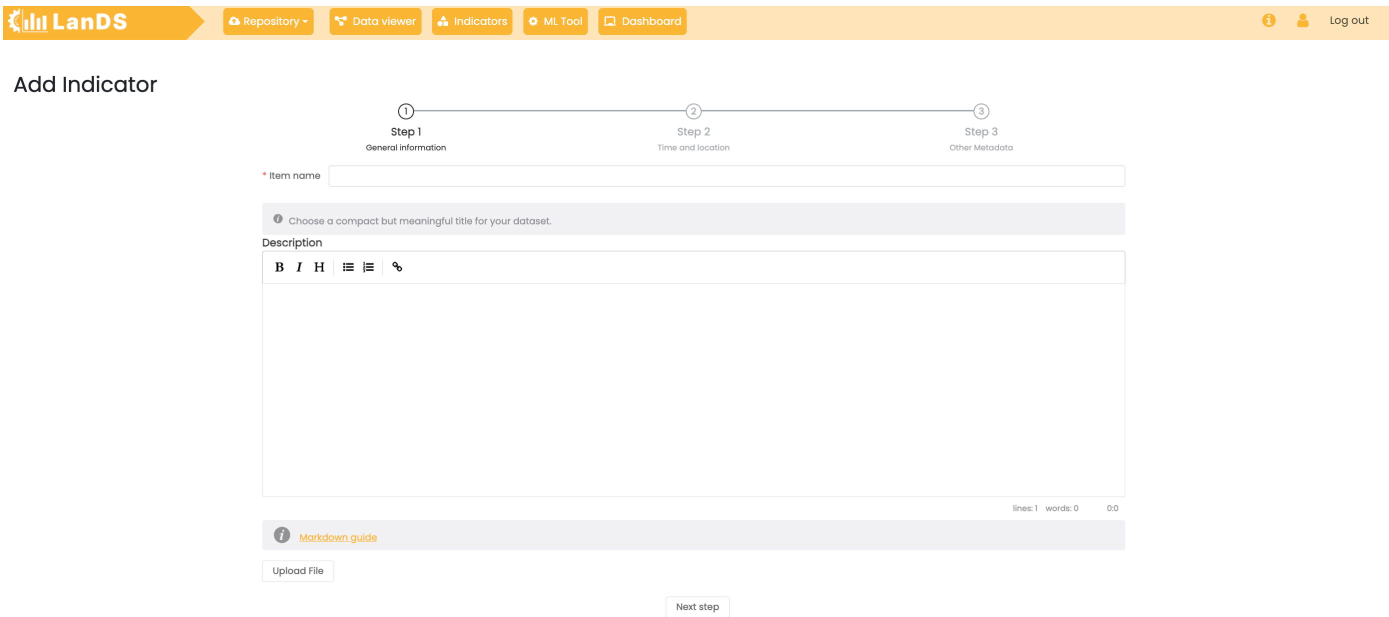


Figure 9 Example of dataset (item) upload form.

[8] See https://schema.datacite.org for the full specification.

2.2 Data Viewer

The LanDS data viewer provides quick and effective visualizations of the different datasets uploaded in the repository. Considering the potential diversity of data contents and formats, the data viewer has been developed considering an adaptive approach based on the following principles:

- a selection of different visualisations is available correspondently to different formats, able to explore the variety of dimensions and versions of each dataset (see Table 1);
- templates are/can be provided to authenticated users, in order to provide guidance on the more effective way to produce and upload data on the toolbox;
- whenever data types not previously managed become used and relevant for LanDS users, the toolbox can be easily extended to provide additional visualizations;
- data that cannot be shown can anyhow be uploaded, downloaded and inspected through the extensible set of metadata offered by the repository.

The data viewer is integrated in the LanDS website[9] and can be accessed:

- from the top bar menu (1 in Figure 10), to directly open the data viewer page from the top bar of the LanDS;
- from the dataset/item page (2 in Figure 10), to contextually show the visualizations available for the specific dataset.

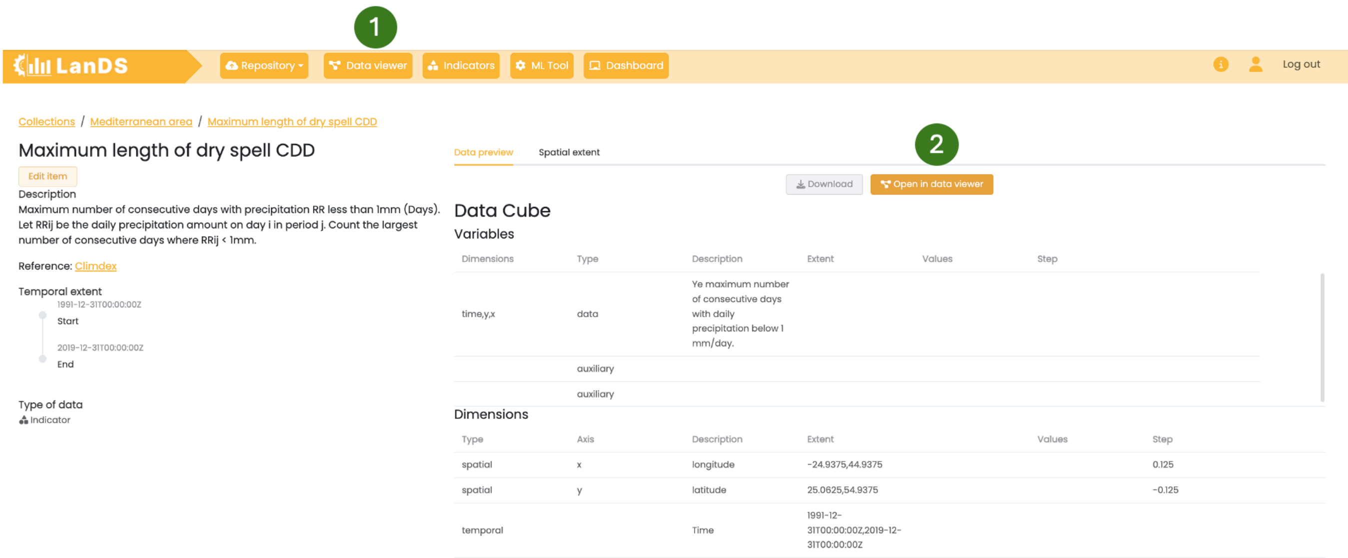


Figure 10 Example of dataset page for the Aridity Index in the Mediterranean Area.

Table 1 Visualization options for different data types.

Dataset type	File formats supported by LanDS	Available chart visualisations on LanDS	Map Visualisations	Technologies applied
Tabular data	.xlsx .csv	Line Monitoring charts Heatmap	N/A	Plotly[10]

[9] http://lands.soft-water.it
[10] https://plotly.com/python/

Dataset type	File formats supported by LanDS	Available chart visualisations on LanDS	Map Visualisations	Technologies applied
Geographical data	.shp (in .zip) .gpkg .kml / .kmz WKT .csv/.txt .tif / .asc	N/A	Geo Map Choropleth Map Tif map with slider to navigate different bands (e.g., different time step)	Leaflet[11] Plotly[2]
Distributed Time series	.nc	Time series line plot for spatially averaged data, with range of reference values Boxplot and decade boxplot for decades period	Raster map with time slider to navigate different time step	Plotly[2] Hvplot[12]

The Data viewer page presents a select box (as can be seen on top of Figure 11 11) that allows to filter and select among available datasets: once selected a dataset, one or more visualization options are presented to the user, depending upon the type of dataset and the associated file, as reported in Table 1and addressed in the next subsection 2.2.1.

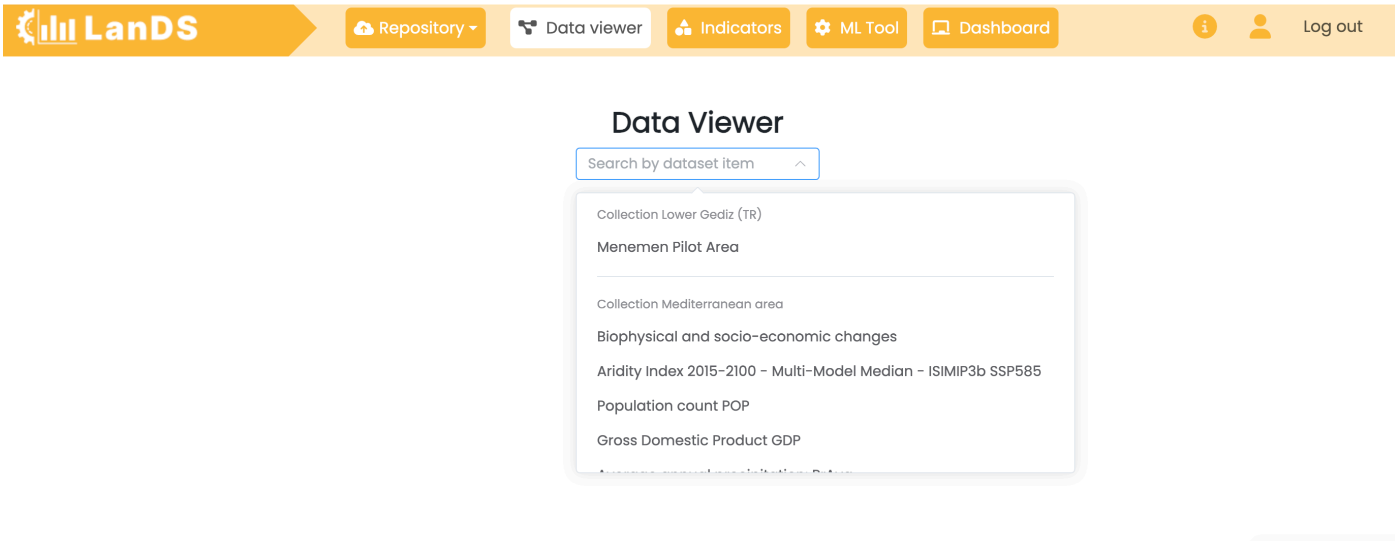


Figure 11 Data viewer page, showing the select box to choose the dataset to be visualized, belonging to a specific collection (the search functionality is set on the dataset item).

2.2.1 Data viewer visualizations

This section presents different types of visualisations embedded in the LanDS data viewer through several examples showcasing them. The following functionalities can be appreciated:

- interactive chart: all charts come with a toolbox allowing to interact with the chart by panning, zooming in-out, switching on/off the single traces, reset axis, hovering the data to inspect single values and, finally, exporting the chart as static image;

[11] https://leafletjs.com/
[12] https://hvplot.holoviz.org/

- multi-panel plot: whenever relevant, the data viewer can split visualisations in two or more panels to present several datasets at once;
- contextual units and title: whenever available and formally described in the dataset, automatic extraction of metadata allows to customise several chart components like axis labels, legend names, units of measurement;
- dimensionality reduction: to allow quick but informative visualisation of complex dataset, a dimensional reduction operation can be performed.

Figure 12 presents an example of time series line plot with the associated reference range related to the *Aridity Index* dataset, part of the *Mediterranean Area* collection, in which time series of yearly values are averaged over their spatial dimension to provide min, mean and max values for each year.

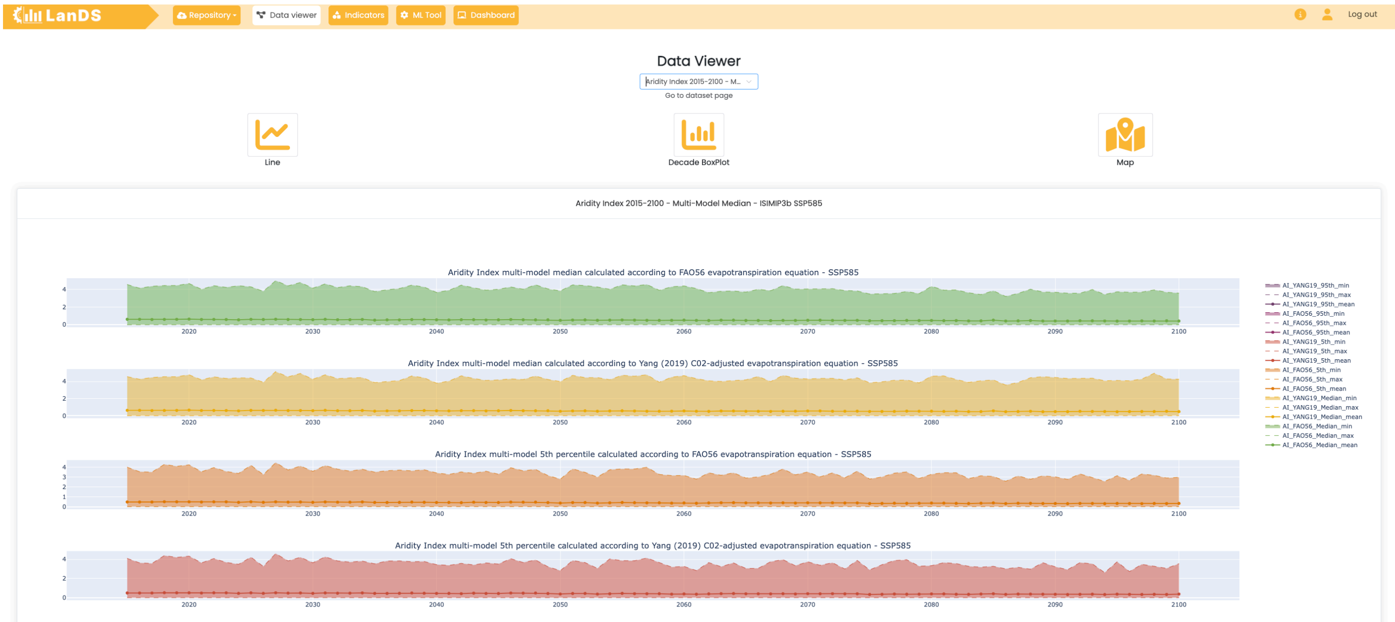


Figure 12 Example of the Aridity Index in the Mediterranean Area in the LanDS data viewer: time series line plot with range reference.

Figure 13 displays the time series *decade boxplot* for the same dataset. In general, the *decade boxplot* visualization can be useful to detect decadic trends in long time series: in order to generate it, the dataset is dynamically divided into decades and box-plot statistics[13] are computed for each decade and for each data variable. In the specific case of the Aridity Index dataset, it presents the results of the aggregation over outputs coming from different models, through 3 statistics: the multi-model median, the multi-model 5th percentile and the multi-model 95th percentile.

[13] Namely: minimum, maximum, median, first and third quartile, Interquartile range

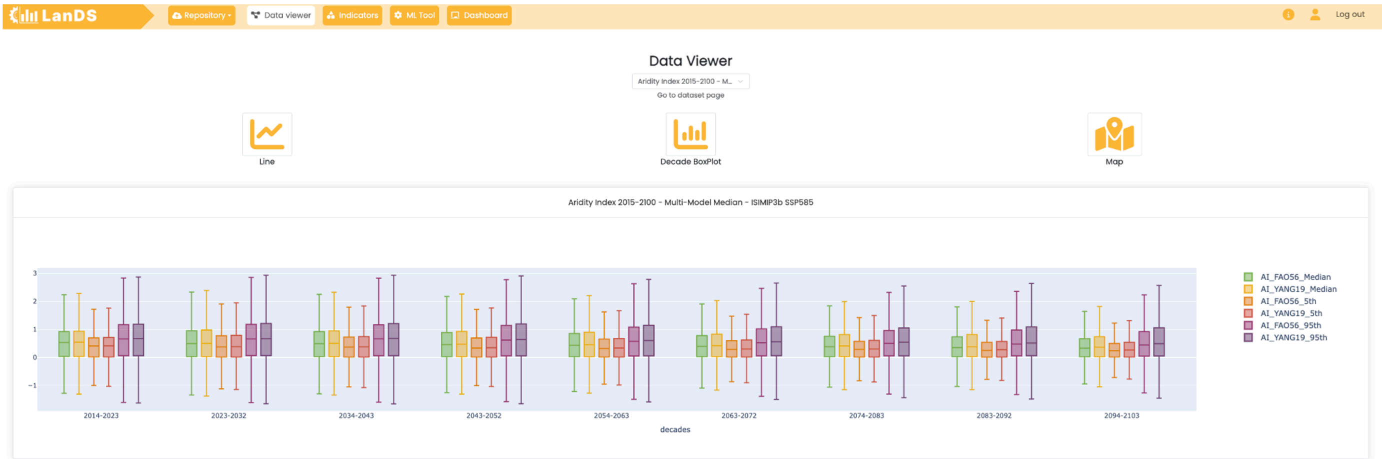


Figure 13 Example of the Aridity Index in the Mediterranean Area in the LanDS data viewer: decade boxplot visualization.

Figure 14 presents an example of simple map visualization for the Menemen Pilot Area dataset, related to the Lower Gediz (Turkey) collection.

The file uploaded is a package (.kmz or .zip) with a unique feature representing the boundary of the case study area and it is represented as overlay with a tile basemap provided by OpenStreetMap[14].

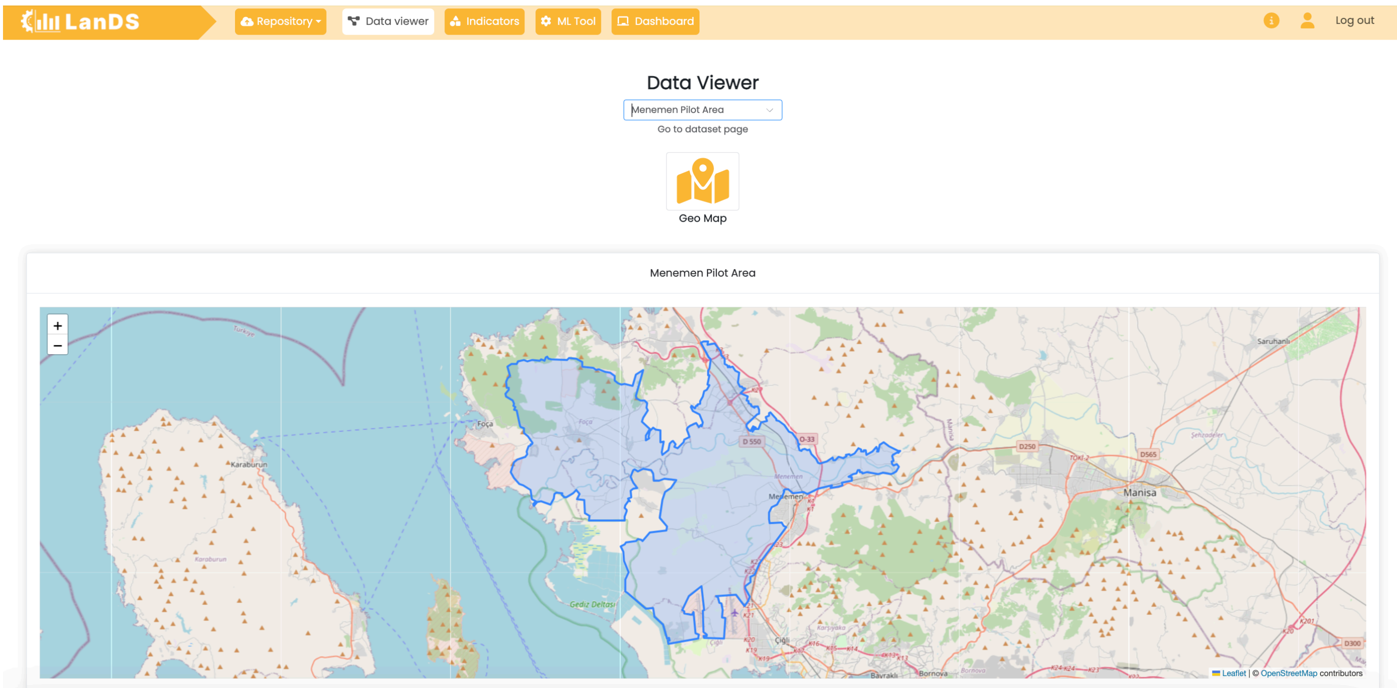


Figure 14 Lower Gediz study area dataset: simple map representation.

Figure 15 15 presents an example of choroplet map (also known as categorised map) for the Soil Texture dataset, related to the Troodos Mountains (Cyprus) collection.

---

14 A free, open geographic database updated and maintained by a community of volunteers via open collaboration (www.openstreetmap.org)

The file uploaded is a shapefile (.shp) with a specific field containing information that can be used to represent the soil texture features of the study area (the shapefile must be uploaded as a zipped folder .zip containing the .shp and the accompanying .shx, .dbf, .cpg files). The map shows the different features in the dataset styled according to the different values present in the categorisation field, overlapped with a tile basemap, always based on OpenStreetMap but with a different visualization style.

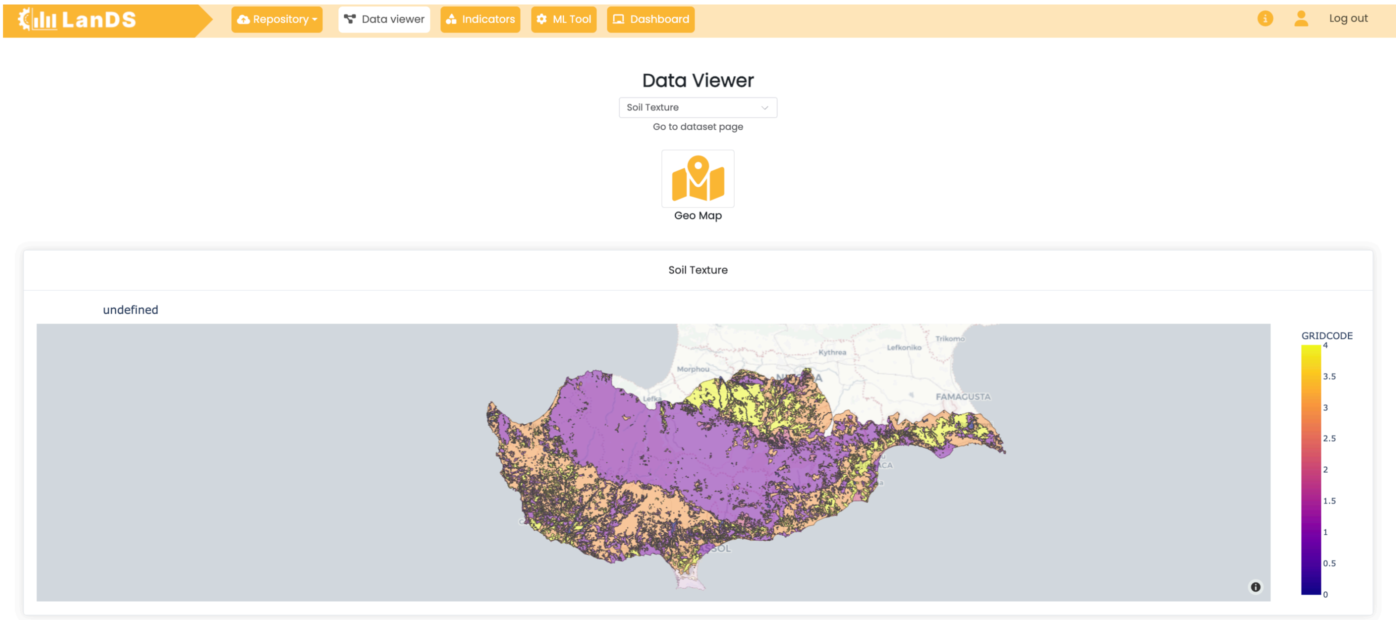


Figure 15 Troodos Mountains (Cyprus) - Soil Texture dataset: Choropleth map.

More visualization options (e.g., monitoring charts) are presented in subsection 2.3.2 and visualization types can be selected by a logged in user in the editing form of a dataset, depending on its format, as shown in Figure 16.

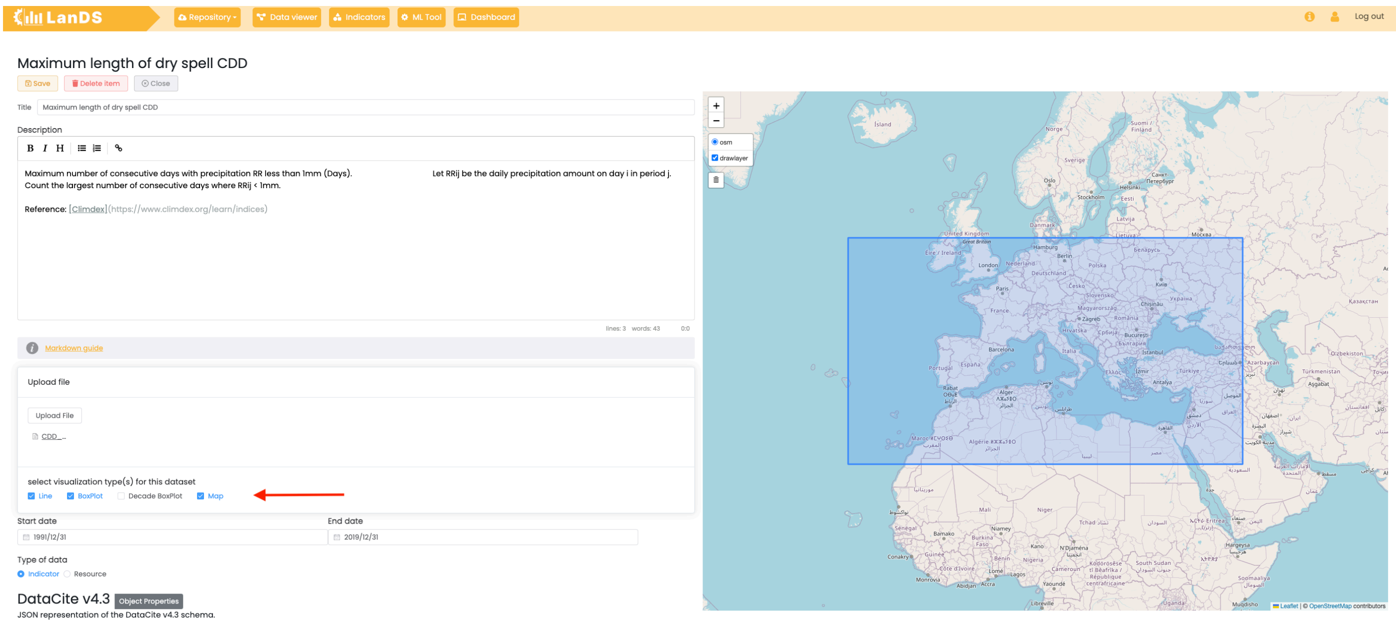


Figure 16 Editing page of a dataset (in this case, the CDD index), in which a logged in user can select the desired visualization type on the LanDS.

2.3 Indicators Library

The Indicators Computational Code Library (or indicator library) is a set of processing routines, that implement a set of indicators both quantifying land degradation processes and land restoration action effects. Indicators are selected from the set of biophysical, climate, and socio-economic indicators identified in WP2 for the Mediterranean area and from WP3 on the local scale in the ERLLs/pilot areas.

The library implements the computation of standardized indicators in the LanDS, aimed to provide quantitative metrics supporting analysis and comparisons across different spatial and temporal scales. For the Mediterranean area, a list of 56 indicators have been implemented and computed starting from biophysical, climate, and socio-economic datasets compiled for the whole Mediterranean area in WP2. These datasets are based on global repositories and sound literature providing gridded data at different spatial and temporal resolutions (e.g., ISIMIP protocol 3[15], ESACCI[16], Yang et al., 2019, Wang and Sun, 2023), covering past and future periods (see Section 2 of D2.4 *LanDS Prototype* for more details).

The code library is structured as a Jupiter notebook, a web environment where is possible to integrate Python code with rich text description and interactive visualization. The code library is available in a software repository hosted on Gitlab[17] and the code is released under the latest version of the GNU General Public License (GPL)[18].

Table 2 reports the full list of the biophysical, climate, and socioeconomic indicators computed in the current release of the LanDS indicators code library, which encompass climate extreme indices listed in the Annex VI of the IPCC AR6 report (IPCC AR6, 2021), land use changes indicators (e.g. one of those listed is the SDG11.3.1[19]), others related to demographic and economic aspects selected from literature (e.g., Calka et al., 2022; Yang et al., 2023), and characteristics on topography (retrieved from HydroSHEDS[20]).

Table 2 List of the 56 drivers/regional indicators computed on the LanDS code library, starting from D2.4 datasets produced by TUC.

Label	Index Name	Units
TXx	Monthly maximum value of daily maximum temperature	°C
TXn	Monthly minimum value of daily maximum temperature	°C
TNn	Monthly minimum value of daily minimum temperature	°C
TNx	Monthly maximum value of daily minimum temperature	°C
TX90p	Number of days when daily maximum temperature is greater than the 90th percentile	Days
TX10p	Number of days when daily maximum temperature is less than the 10th percentile	Days
TN90p	Number of days when daily minimum temperature is greater than the 90th percentile	Days
TN10p	Number of days when daily minimum temperature is less than the 10th percentile	Days
ID	Number of icing days: annual count of days when TX (daily maximum temperature) <0°C	Days
FD	Number of frost days: annual count of days when TN (daily minimum temperature) <0°C	Days
WSDI	Warm spell duration index: annual count of days with at least six consecutive days when TX >90th percentile	Days

[15] https://www.isimip.org/protocol/3/
[16] https://www.esa-landcover-cci.org/
[17] https://gitlab.com/lands-r4m/indicators_library
[18] https://www.gnu.org/licenses/gpl-3.0.html
[19] https://sdgs.un.org/goals/goal11#targets_and_indicators
[20] https://www.hydrosheds.org/

Label	Index Name	Units
CSDI	Cold spell duration index: annual count of days with at least six consecutive days when TN <10th percentile	Days
SU	Number of summer days: annual count of days when TX (daily maximum temperature) >25°C	Days
TR	Number of tropical nights: annual count of days when TN (daily minimum temperature) >20°C	Days
DTR	Daily temperature range: monthly mean difference between TX and TN	°C
GSL	Growing season length: annual (1 Jan to 31 Dec in Northern Hemisphere (NH), 1 July to 30 June in Southern Hemisphere (SH)) count between first span of at least six days with daily mean temperature TG >5°C and first span after July 1 (Jan 1 in SH) of six days with TG <5°C	Days
Rx1day	Maximum one-day precipitation	mm
Rx5day	Maximum five-day precipitation	mm
R5mm	Annual count of days when precipitation is greater than or equal to 5 mm	Days
R10mm	Annual count of days when precipitation is greater than or equal to 10 mm	Days
R20mm	Annual count of days when precipitation is greater than or equal to 20 mm	Days
R50mm	Annual count of days when precipitation is greater than or equal to 50 mm	Days
CDD	Maximum number of consecutive days with less than 1 mm of precipitation per day	Days
CWD	Maximum number of consecutive days with more than or equal to 1 mm of precipitation per day	Days
R95p	Annual total precipitation when the daily precipitation exceeds the 95th percentile of the wet-day (>1 mm) precipitation	mm
R99p	Annual precipitation amount when the daily precipitation exceeds the 99th percentile of the wet-day precipitation	mm
SDII	Simple precipitation intensity index	Mm/day
SPI6_wet	Standardized precipitation index 6month: yearly cumulated value of monthly index when SPI6 > 1	-
SPI12_wet	Standardized precipitation index 12month: yearly cumulated value of monthly index when SPI12 > 1	-
SPI6_dry	Standardized precipitation index 6month: yearly cumulated value of monthly index when SPI6 < -1	-
SPI12_dry	Standardized precipitation index 12month: yearly cumulated value of monthly index when SPI12 < -1	-
AI_mean	Aridity index FAO: mean over time horizon	-
AI_max	Aridity index FAO: max over time horizon	-
AI_min	Aridity index FAO: min over time horizon	-
SS_mean	Soil salinity: mean over time horizon	ece
SS_max	Soil salinity: max over time horizon	ece
SS_min	Soil salinity: min over time horizon	ece
SFCW	wind speed: mean over time horizon	m/s
PrAvg	Daily average precipitation	mm
PGR	Grow rate at which population size changes during a period. In the annual formulation, it is computed with respect to the first available year of data	[-]
PCR	Change of population (people) in time. In the annual formulation, it is computed each year with respect to the previous one.	%
GCR	Change in gross domestic product converted to international dollars using purchasing power parity rates (PPP) in time	%
LCCR_Ag	Change of land use cover (cropland) in time	%
LCCR_Fo	Change of land use cover (forest land) in time	%
LCCR_Sh	Change of land use cover (schrubland) in time	%

Label	Index Name	Units
LCCR_Se	Change of land use cover (settlement) in time	%
LCNSR	Land consumption rate	-
LCR_PGR	Ratio of land consumption rate (settlement/urban build-up area growth rate) to population growth rate	-
POP	Populatione mean over the time horizon	persons
GDP	GDP mean over the time horizon	USD PPP
LCAg	Land cover Agriculture: mean over the time horizon	%
LCFo	Land cover Forestry: mean over the time horizon	%
LCSe	Land cover Settlement: mean over the time horizon	%
LCSh	Land cover Shrubland: mean over the time horizon	%
DEM	Elevation above mean sea level	m
Slope	Slope	%

The computed indicators are uploaded in the repository and visualised in the data viewer, as presented in some examples shown later in subsection 2.3.1.

2.3.1 Indicators on the LanDS

Indicators outcomes[21] are stored in the data repository (see Section 2.1) and can be visualized in the data viewer (described in Section 2.2). They will be integrated in the Web Dashboard as well (D4.4 *LanDS Dashboard* due in M34), to support analysis and comparison of project outcomes across different PAs, assess possible future scenarios and extend the analysis up to the Mediterranean regional scale. In fact, the indicators on the LanDS can be used to support the assessment of implemented restoration actions in WP5 and to orient policy recommendations in WP6.

This section is intended to show exemplary visualizations of the indicators on the LanDS.

As shown in Figure 7, the data browser displays all data as list of datasets and files, which are indicated with a title and classified by type (resource or indicator), useful also for filtering options. By clicking on a specific indicator name, the dataset page is accessed with metadata related to the indicator, like spatial and temporal coverage, units, and description (Figure 8). From this page, it is possible to both download the dataset of computed indicators or access to available data visualization in the data viewer section.

Figure 17, Figure 18 and Figure 19 report examples of a climate extreme indicator (stored as NetCDF file in the repository) available visualizations in the data viewer (respectively, line plot, box plot, and map with time slider), while Figure 20 and Figure 21 report a couple of examples of a satellite-based indicators, which are typically uploaded as text, TIFF, NetCDF or Shape files (respectively, .csv, .tif, .nc, .zip or .kmz including .shp), and how they can be visualized in the LanDS.

[21] In general, any indicators can be uploaded and visualized in the LanDS (not only the indicators computed by the LanDS code library).

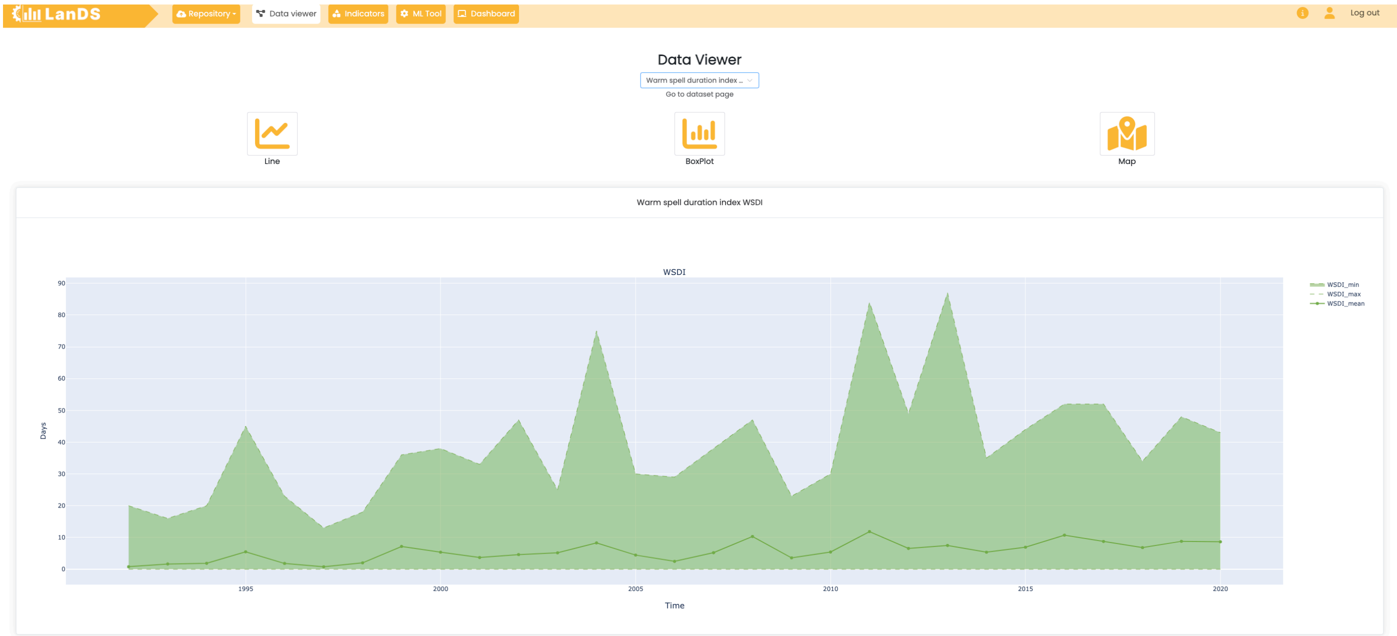


Figure 17 Climate extreme indicator WSDI (NetCDF) line plot visualization in the LanDS: in this case, WSDI is shown.

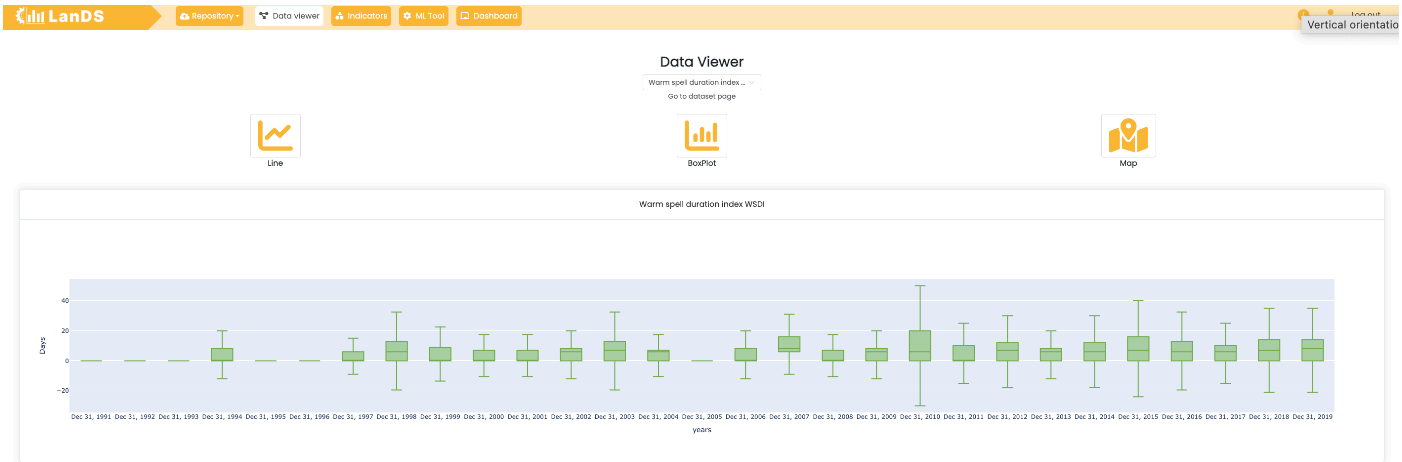


Figure 18 Climate extreme indicator (NetCDF) boxplot visualization in the LanDS: in this case, WSDI is shown.

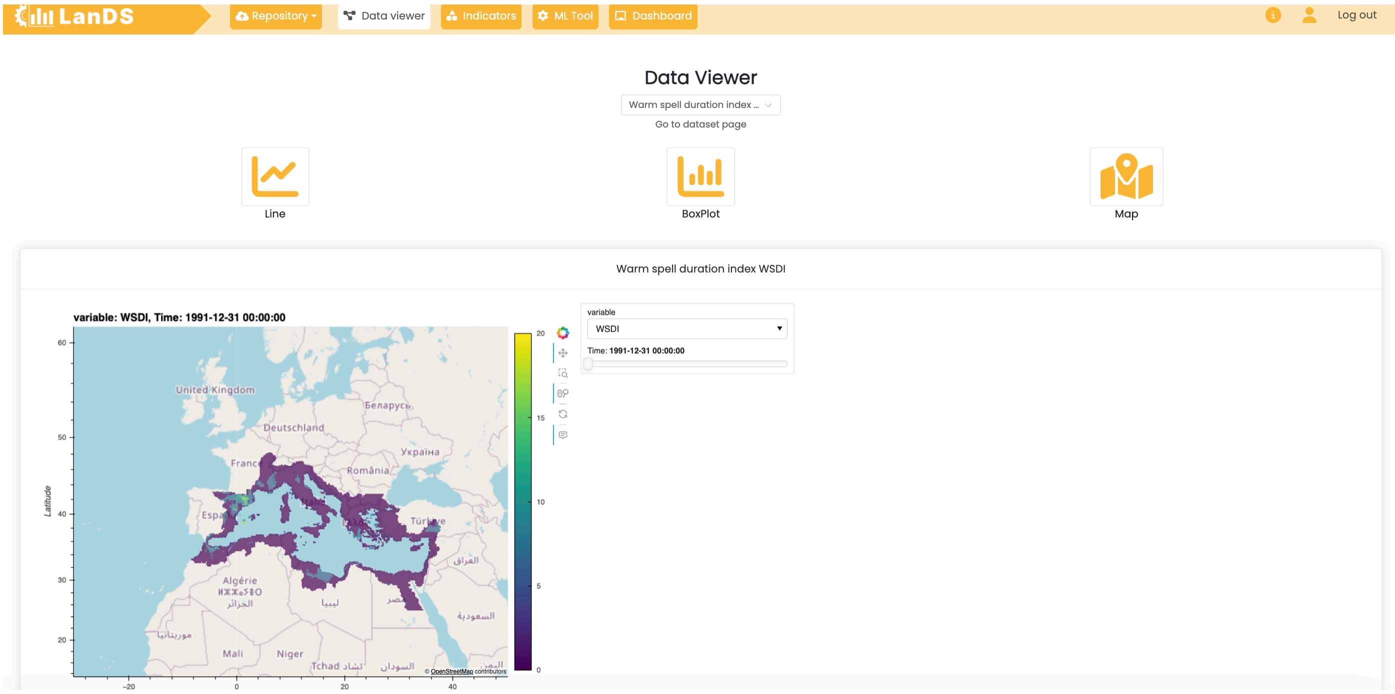


Figure 19 Climate extreme indicator (NetCDF) map visualization in the LanDS: in this case, WSDI is shown. A time slider is available on the right-hand side of the map to select the desired year to be displayed.

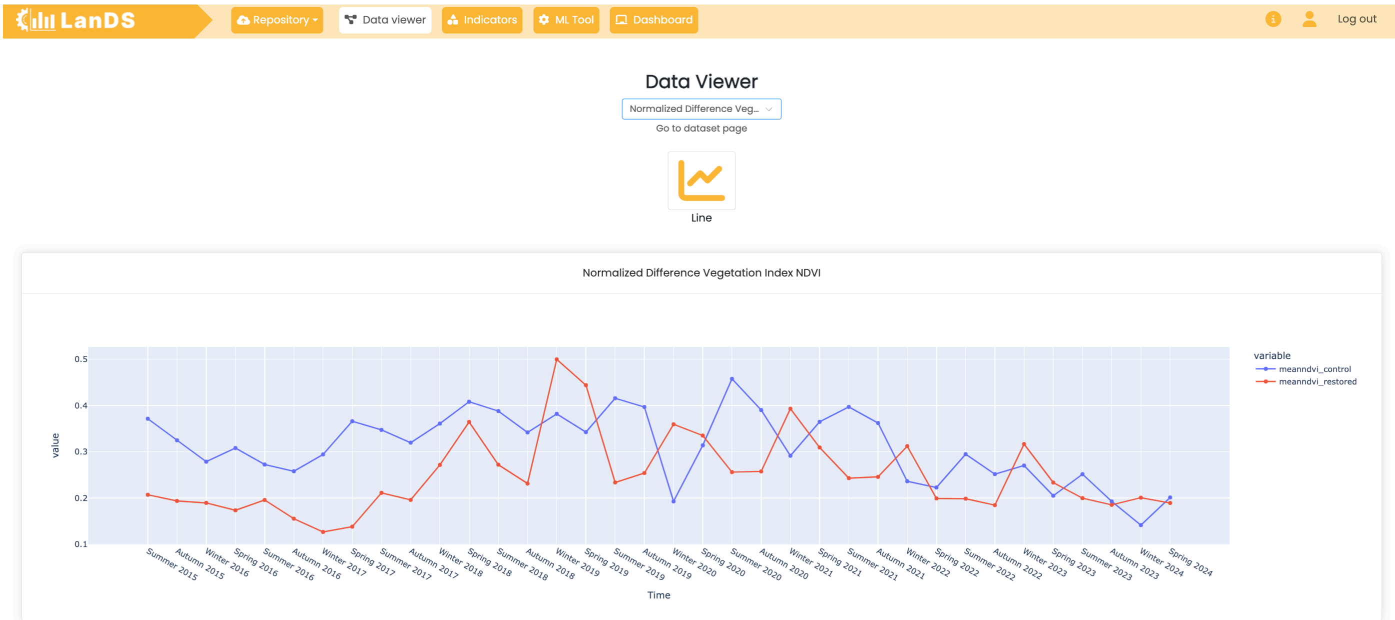


Figure 20 Satellite-based indicator (NDVI: Normalized Difference Vegetation Index) uploaded as Text file (.csv) to compare control and restored areas in Trodoos Mountains collection.

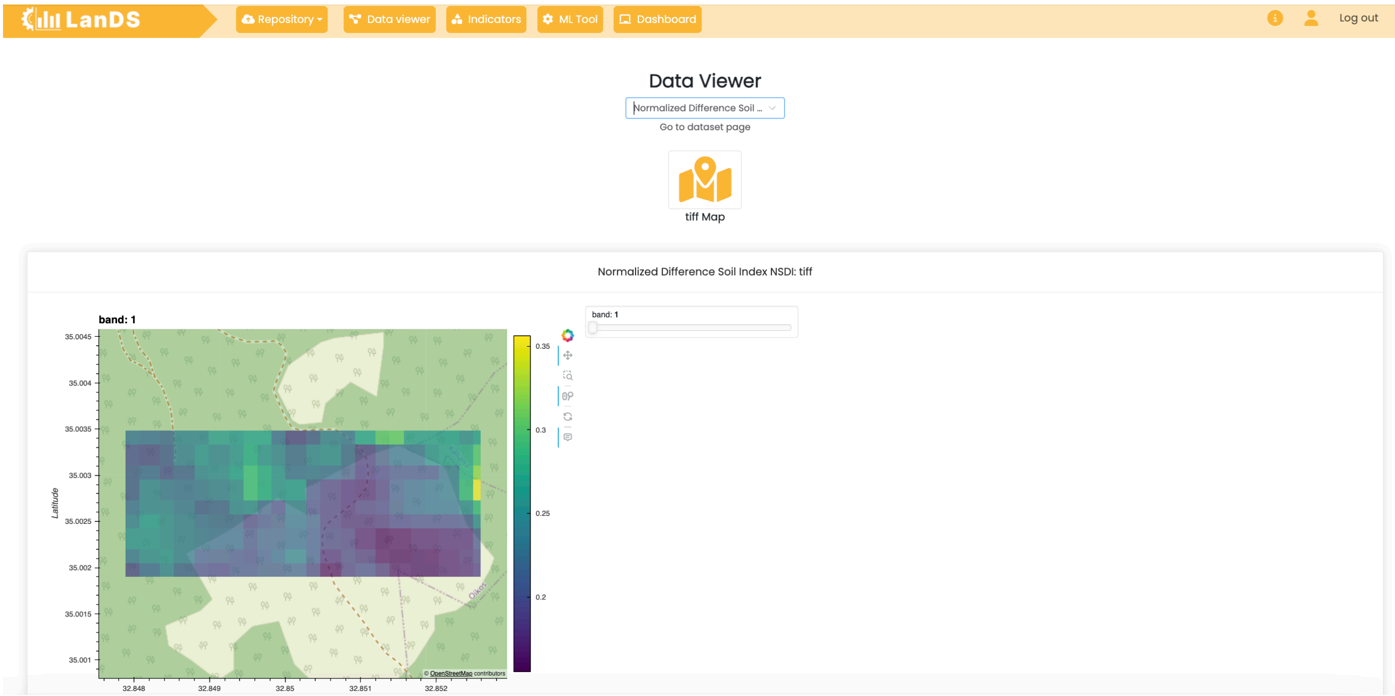


Figure 21 Satellite-based indicator (NDSI: Normalized Difference Soil Index) uploaded as TIFF file (.csv) to monitor the restoration action in Trodoos Mountains collection.

2.3.2 Monitoring data as example of local scale indicators

At the local scale, sampling campaigns are periodically run in most of the PAs to monitor the effects of the implementation of restoration actions (e.g., comparing sites where a restoration action is implemented with others where it is not). Correspondent data can be stored in the repository as indicators[22] and eventually visualized in the data viewer, if compliant with the templates and visualization types supported by the LanDS (Table 1).

To support WP5 in setting up a monitoring system for Task 5.4, a *monitoring data template* (Table 3) was created, ready to receive required information: the parameters measured, the type of experiment/action implemented, and the different sites in which the samples are taken.

Table 3 Monitoring data template for effective monitoring of implemented restoration actions.

Indicator	Units	Reference time (MM/DD/YYYY)	Experiment 1 Site 1	Experiment 1 Site 2	…	Experiment 2 Site 1	Experiment 2 Site 2	…
e.g., Total Nitrogen	e.g., mm/kg	Sampling campaign date (e.g., 11/24/2023)	…	…	…	…	…	…
e.g., Organic matter	e.g., t/ha	..	…	…	…	…	…	…
…	…	…	…	…	…	…	…	…
as many as needed	…							

The monitoring data file with real measurements for Stornara and Tara collection (PA3) is already uploaded to the repository and can be visualized in the data viewer (Figure 22). The correspondent sampling points are

[22] Monitoring data are conceived as *indicators* as they assess the effectiveness of a specific restoration action.

also uploaded as a GeoPackage file (.gpkg) and can be visualized as geo map as shown in Figure 23, including information on latitude, longitude, type of restoration action, and soil depth.

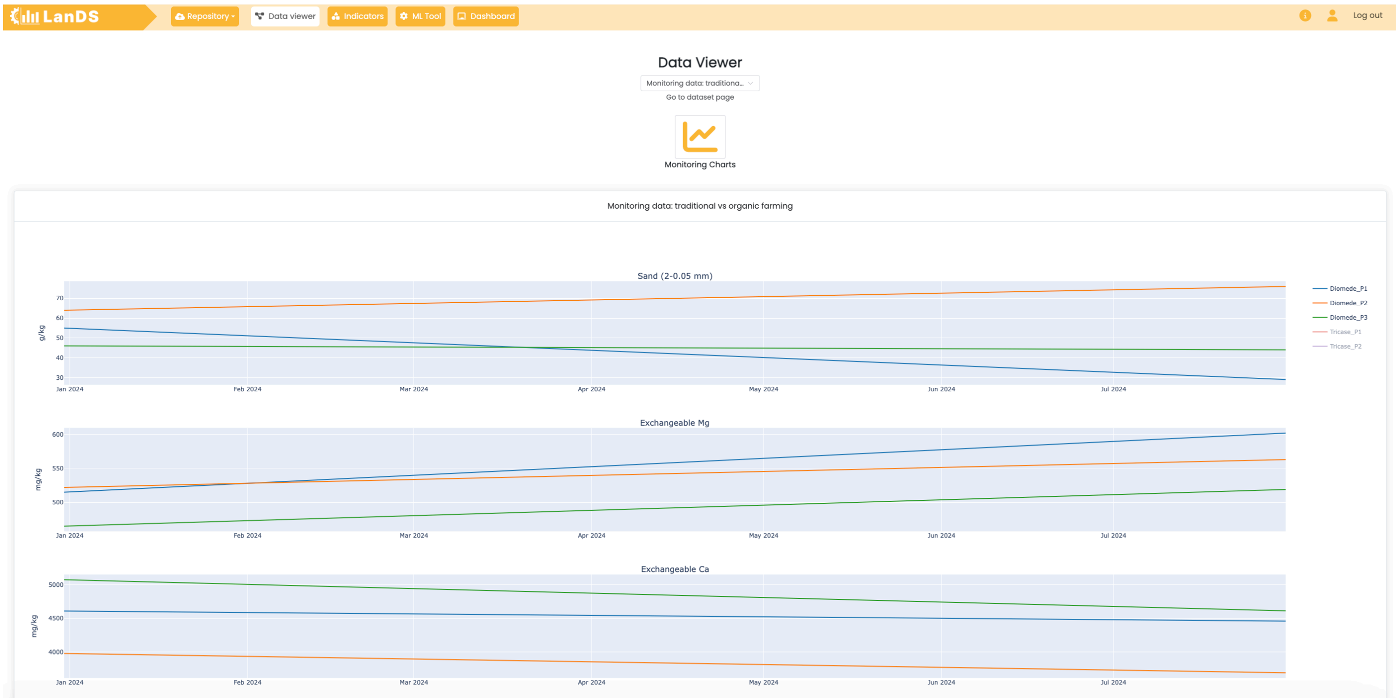


Figure 22 Monitoring charts resulting from the two sampling campaigns run in Stornara and Tara, comparing traditional farming (Tricase in two sites, hidden in the charts) and organic farming (Diomede in three sites, shown in the charts).

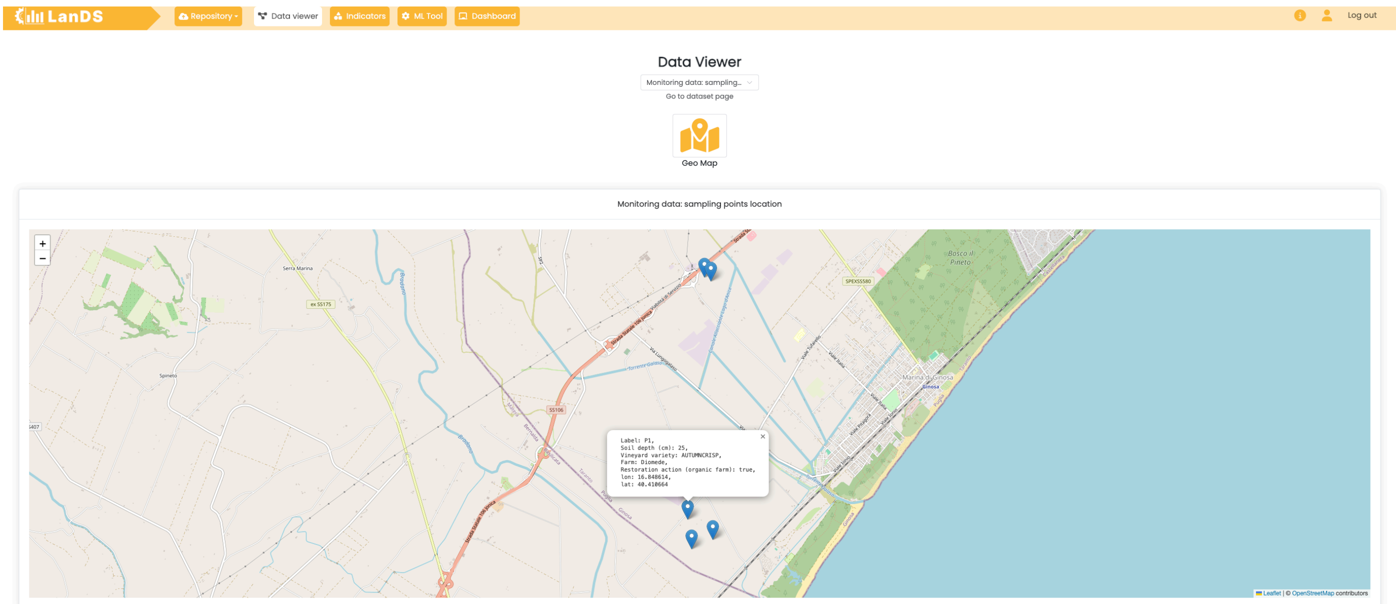


Figure 23 Monitoring data sampling points for Stornara and Tara PA, including information on latitude, longitude, and soil depth.

2.4 Machine Learning Procedure

The LanDS ML procedure combines available global and regional datasets with site specific characteristics, leveraging on the knowledge and expertise coming from ERLLs, to classify different areas in terms of climatic trends and socio-economic circumstances, land degradation severity and applicability for the restoration measures considered in the project. The ML tool can be sourced with historical and future climate and socio-economic scenarios, based on bio-physical, climate and socioeconomic variables (computed from WP2 data), to support better-informed land restoration actions and impact assessment (WP5) and to provide useful insights for policy recommendations and support decision-making processes (WP6).

2.4.1 Inputs

The different spatial and temporal resolutions and time horizons of the original datasets (see Section 2 of D2.4 *Past and Future Drivers of Change*) led to a necessary compromise to create a consistent input dataset for the ML procedure, which is consequently characterised as follows:

- Temporal coverage: 2001 – 2019, the only time horizon common for all data;
- Temporal resolution: all indicators have been computed at a yearly time step and later aggregated over the time horizon or every five years, depending on the configuration run;
- Spatial extent: Mediterranean region, defined by the hydrological basins draining towards the Mediterranean Sea from Hydrobasins 05 level (Lehner and Grill, 2013), as used also in the specific climate classifications for the Mediterranean (Allam et al., 2020) and adding the subbasins belonging to Merchouch (MO) and Tamia (EG), to include REACT4MED Pilot Areas in Morocco and Egypt, respectively;
- Spatial resolution: 0.125 x 0.125 degrees.

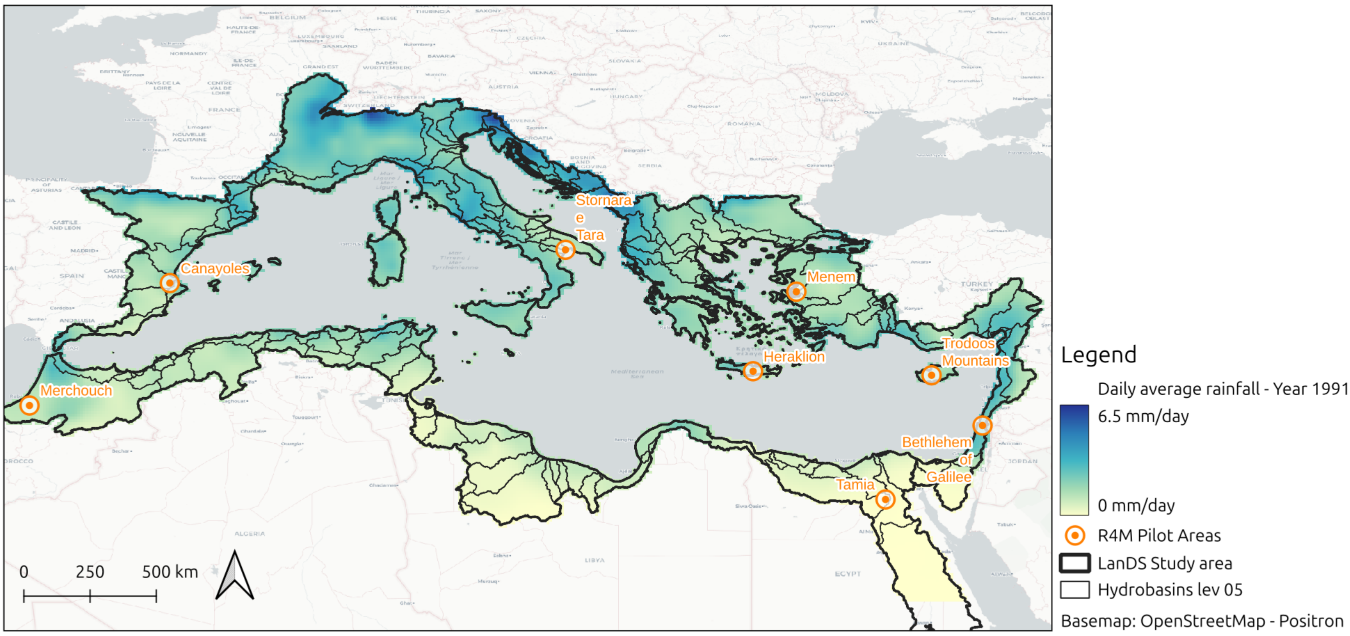


Figure 24 Precipitation average index (PrAvg) averaged for year 1991, covering the Mediterranean region considered for the ML procedure.

An example of input is shown in Figure 24 displaying the Precipitation Average Index (PrAvg) for year 1991, at the spatial extent considered for the ML tool.

2.4.2 Methodology

The LanDS ML procedure for the identification of potentially suitable areas in the Mediterranean to up- and outscale REACT4MED restoration actions can be summarised in the following working steps (schematically presented in Figure 25):

1. *The candidate ML tool inputs are computed for a reference past period, using the LanDS-computed indicators based on Task 2.4 datasets;*
2. *Input Variable Selection (IVS) is run checking the candidate inputs relevance and redundancy by a correlation analysis and a Principal Component Analysis (PCA), to eventually reduce the inputs set to the most significant ones;*
3. *A clustering analysis (k-means) based on the inputs set obtained in step 2 is run to find areas with similar characteristics in the Mediterranean region.*

The outcomes of step 3 will serve the expert-based analysis (see subsection 2.4.4) integrating higher resolution datasets and specific local conditions to find potentially suitable areas for each specific project restoration action.

The software code has been released on the LanDS Gitlab repository[23] and can be used to run step 2 and 3 of the procedure.

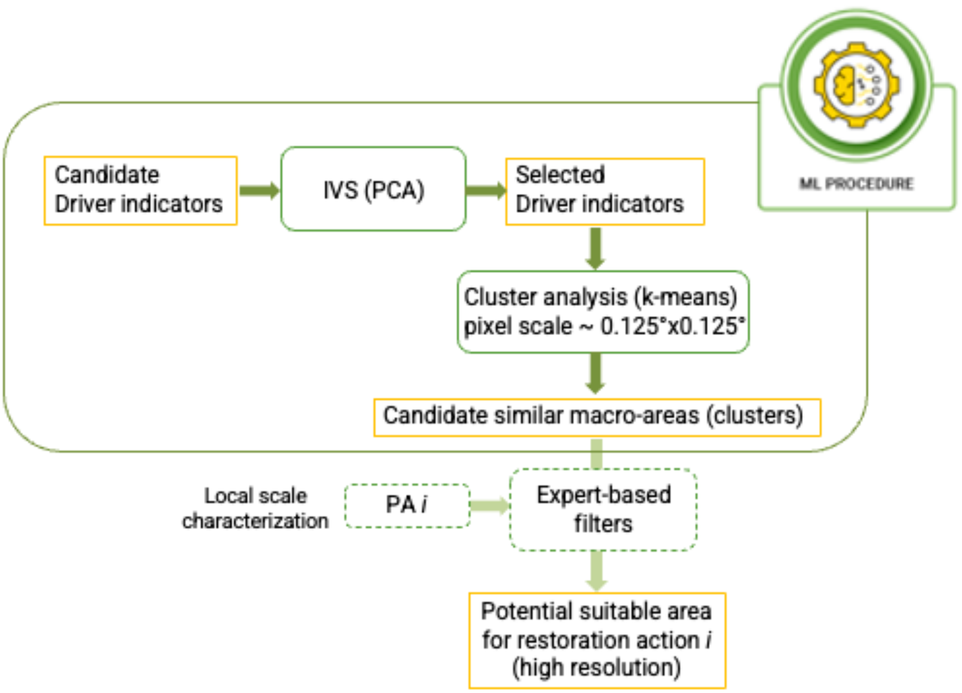


Figure 25 ML procedure workflow for the historical scenario (2001-2019).

In the following, we describe the implemented methodological steps and the choices taken for the reported application case (subsection 2.4.3), while other runs and correspondent configurations are listed/shown in the Annex.

2.4.2.1 Input Variable Selection

The set of candidate inputs has been intentionally created to include an exhaustive description of biophysical and socio-economic characteristics of the study area; nevertheless, some of them are potentially redundant

[23] https://gitlab.com/lands-r4m/ml_tool

or not relevant. For this reason, following a data-driven approach, we run a preliminary input selection process to choose the ML tool input set starting from the initial list of Table 2, by applying:

- A-priori selection of indicators based on their spatial distribution and temporal variability, comparing historical and future scenarios datasets (from 56 to 40 indicators, reported in Table 4);
- A correlation analysis, a statistical method used in research to measure the strength of the linear relationship between two variables and compute their association. Looking at Figure 26, the set of biophysical and climate indicators is clearly identified by the sub-matrix with higher absolute values of correlation, while those related to socio-economic values present lower values. Starting from these results, we applied a threshold to iteratively exclude all indicators providing identical information (i.e., highly correlated) to others already in the set. Correlation is run within the categories reported in Table 4, defined to bind significant indicators (e.g., Population) and avoiding their exclusion from the input dataset during this step;
- A Principal Component Analysis (PCA), a widely applied statistical technique to decrease the dimensionality of datasets while preserving the most representative and uncorrelated variables (Jolliffe, 2002; Allam et al., 2020). It can provide indirect information on the relevance of the input features and this information is exploited to eventually exclude non relevant information.

Table 4 List of the 40 ML tool inputs, along with their short definition, unit of measure, and belonging category. The 40 ML inputs are selected from the longer list of the 56 climate, biophysical and socioeconomic indicators computed by the LanDS (reported in Table 2).

Label	Index Name	Units	Category
TXx	Monthly maximum value of daily maximum temperature	°C	Min and Max temperature
TXn	Monthly minimum value of daily maximum temperature	°C	
TNn	Monthly minimum value of daily minimum temperature	°C	
TNx	Monthly maximum value of daily minimum temperature	°C	
ID	Number of icing days: annual count of days when TX (daily maximum temperature) <0°C	Days	
FD	Number of frost days: annual count of days when TN (daily minimum temperature) <0°C	Days	
WSDI	Warm spell duration index: annual count of days with at least six consecutive days when TX >90th percentile	Days	
CSDI	Cold spell duration index: annual count of days with at least six consecutive days when TN <10th percentile	Days	
SU	Number of summer days: annual count of days when TX (daily maximum temperature) >25°C	Days	
TR	Number of tropical nights: annual count of days when TN (daily minimum temperature) >20°C	Days	
DTR	Daily temperature range: monthly mean difference between TX and TN	°C	Mean Temperature
GSL	Growing season length: annual (1 Jan to 31 Dec in Northern Hemisphere (NH), 1 July to 30 June in Southern Hemisphere (SH)) count between first span of at least six days with daily mean temperature TG >5°C and first span after July 1 (Jan 1 in SH) of six days with TG <5°C	Days	
Rx1day	Maximum one-day precipitation	mm	High Precipitation
Rx5day	Maximum five-day precipitation	mm	
R5mm	Annual count of days when precipitation is greater than or equal to 5 mm	Days	

Label	Index Name	Units	Category
R10mm	Annual count of days when precipitation is greater than or equal to 10 mm	Days	
CDD	Maximum number of consecutive days with less than 1 mm of precipitation per day	Days	
CWD	Maximum number of consecutive days with more than or equal to 1 mm of precipitation per day	Days	
R95p	Annual total precipitation when the daily precipitation exceeds the 95th percentile of the wet-day (>1 mm) precipitation	mm	
R99p	Annual precipitation amount when the daily precipitation exceeds the 99th percentile of the wet-day precipitation	mm	
SDII	Simple precipitation intensity index	mm day−1	Mean Precipitation
PrAvg	Daily average precipitation	mm	
SPI6_wet	Standardized precipitation index 6month: yearly cumulated value of monthly index when SPI6 > 1	-	SPI_wet
SPI12_wet	Standardized precipitation index 12month: yearly cumulated value of monthly index when SPI12 > 1		
SPI6_dry	Standardized precipitation index 6month: yearly cumulated value of monthly index when SPI6 < -1	-	SPI_dry
SPI12_dry	Standardized precipitation index 12month: yearly cumulated value of monthly index when SPI12 < -1		
LCAg	Land cover Agriculture: mean over the time horizon	%	Land cover
LCFo	Land cover Forestry: mean over the time horizon	%	
LCSe	Land cover Settlement: mean over the time horizon	%	
GDP	GDP mean over the time horizon	USD PPP	Economic Indicator
POP	Populatione mean over the time horizon	persons	Social Indicator
SFCW	wind speed: mean over time horizon	ece	Wind
AI_mean	Aridity index FAO: mean over time horizon	-	Land degradation indicators
AI_max	Aridity index FAO: max over time horizon	-	
AI_min	Aridity index FAO: min over time horizon	-	
SS_mean	Soil salinity: mean over time horizon	ece	
SS_max	Soil salinity: max over time horizon	ece	
SS_min	Soil salinity: min over time horizon	ece	
DEM	Elevation above mean sea level	m	Topography
Slope	Slope	%	

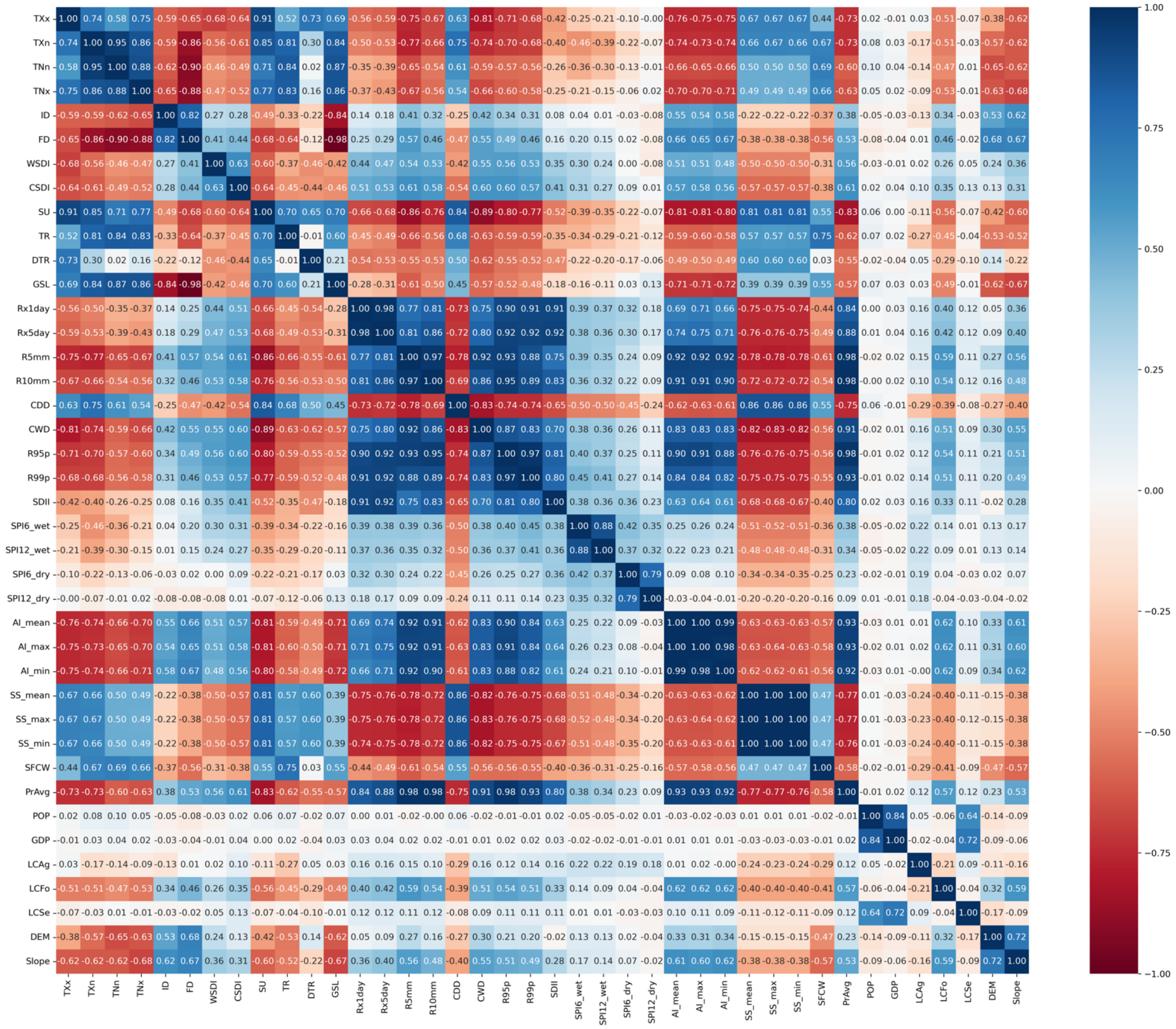


Figure 26 Original correlation matrix, based on the 40 indicators listed in Table 4 (selected from the original 56 LanDS indicators).

2.4.2.2 Clustering analysis with k-means

Once having determined the input dataset, we run a clustering analysis to classify the different areas in the Mediterranean region into a certain number of clusters with similar characteristics.

Clustering analysis is a technique in data mining and statistics used to group a set of objects in such a way that objects in the same group (called cluster) are more similar to each other than to those in other groups (clusters). The main goal of clustering is to identify and categorize data into meaningful subgroups without prior knowledge of the group definitions. In this case, we deploy one of the most popular clustering methods: the k-means method introduced by Edward Forgy (Forgy, 1965) and MacQueen (1967). The goal of k-means is to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster; thus, minimizing the square error objective function for distance optimization. One limitation of k-means is that the coefficient k (number of clusters) must be specified in advance. Nevertheless, there are different approaches that support the proper choice of k, such as the Silhouette Analysis, which provides a measure of how similar an object is to its own cluster compared to other clusters.

We run k-means clustering analysis considering different values of k (from 3 to 15) and look at the correspondent silhouette plots. Each plot displays the silhouette coefficient for each data point, ordered by their cluster assignment. Then, we compare the average silhouette scores for different values of k. The value of k with the highest average silhouette score is typically considered the best choice for the number of clusters, but additional considerations can also support the choice.

2.4.3 Application

The application case of the ML tool that we decided to report is characterised by a candidate input dataset composed of a matrix of 56 indicators (columns), each of them averaged over the 19 years-time horizon (2001-2019). These are reduced to 40 indicators (Figure 26) when the a-priori selection is applied. Then, we run a correlation analysis to eventually filter out redundant information. After several test runs varying the correlation threshold, we decided to set it equal to 0.70 and obtaining a reduced dataset of 20 indicators (Figure 27).

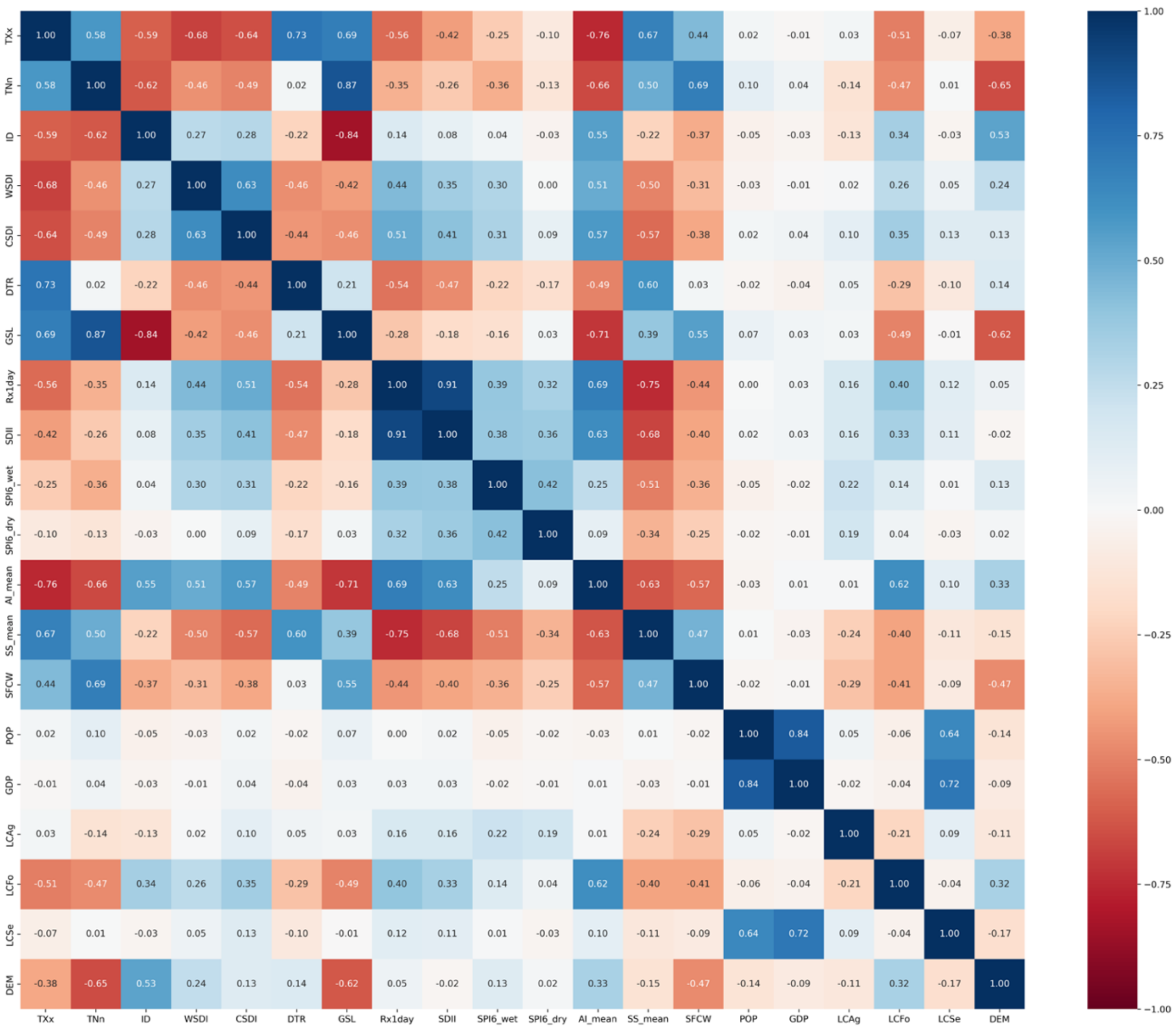


Figure 27 Reduced correlation matrix applying a threshold of 0.70 and resulting into a heatmap with 20 indicators.

Subsequently to the correlation analysis, we run the PCA. Figure 28 reports the PCA variance plot: the y-axis of the plot represents the eigenvalues of the covariance matrix, which correspond to the amount of variance explained by each PC, while the x-axis represents the PCs in order of importance. The cumulative ratio of total variance explained by the PCs (in red) helps in understanding how many components are needed to get the desired amount of explained variance. As first tentative value we considered the PC that provide the 90%

of explained variance (green dashed line, reached with 10 PC) and checked if there were inputs that were not relevant for any of the 10 PCs.

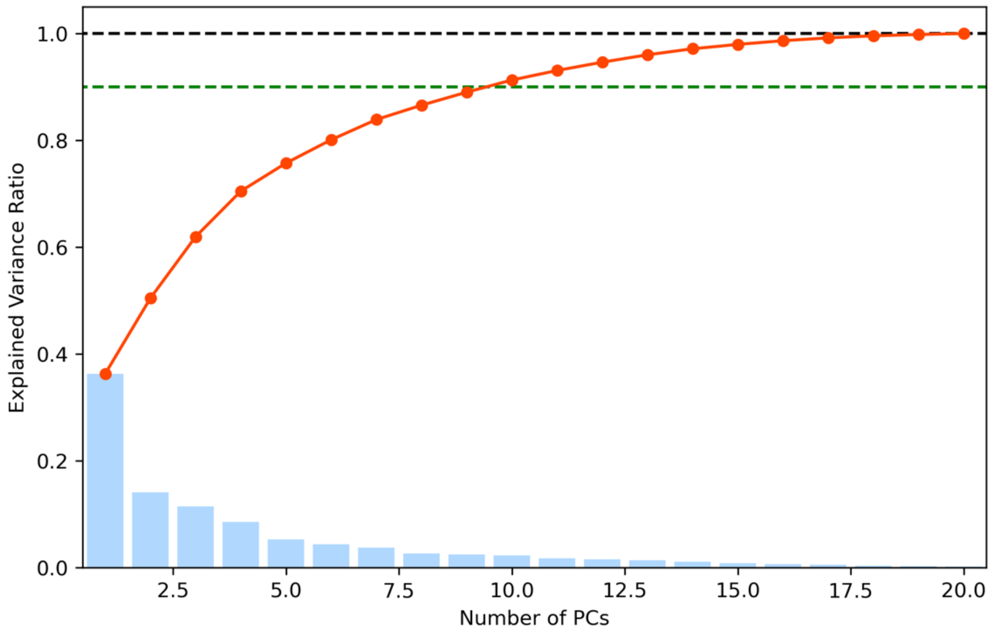


Figure 28 PCA variance plot, where the variance is represented by the red curve: the black dashed line indicates the threshold of 100% of variance explained (almost reached with 20 PCs), while the green dashed line the 90% (reached with 10 PCs).

Figure 29 reports a heatmap showing the weights of input indicators (on the rows) for the 10 Principal Components selected (columns). Using this kind of visualization, a non-relevant input could be identified with the corresponding row presenting low values (close to 0) across all the columns (PC). Looking at the outcomes of this specific application of the PCA, we decided to keep all 20 indicators as input dataset of the cluster analysis not to lose any important information for a significant/more complete representation of the dataset.

Finally, the k-means clustering analysis is run as third step of the ML procedure. Exploring the outputs resulting from different values of k (from 3 to 15), we chose k=12 as we believe it represents a good compromise as both non-trivial and interpretable result. Figure 30 shows the correspondent output of the clustering analysis and how the twelve clusters are geographically distributed in the Mediterranean region considered for the analysis. We can observe that there is climatically driven classification of the areas but also land use and socioeconomic characteristics are emerging (e.g., demographic concentration in big cities, small islands in a separate cluster).

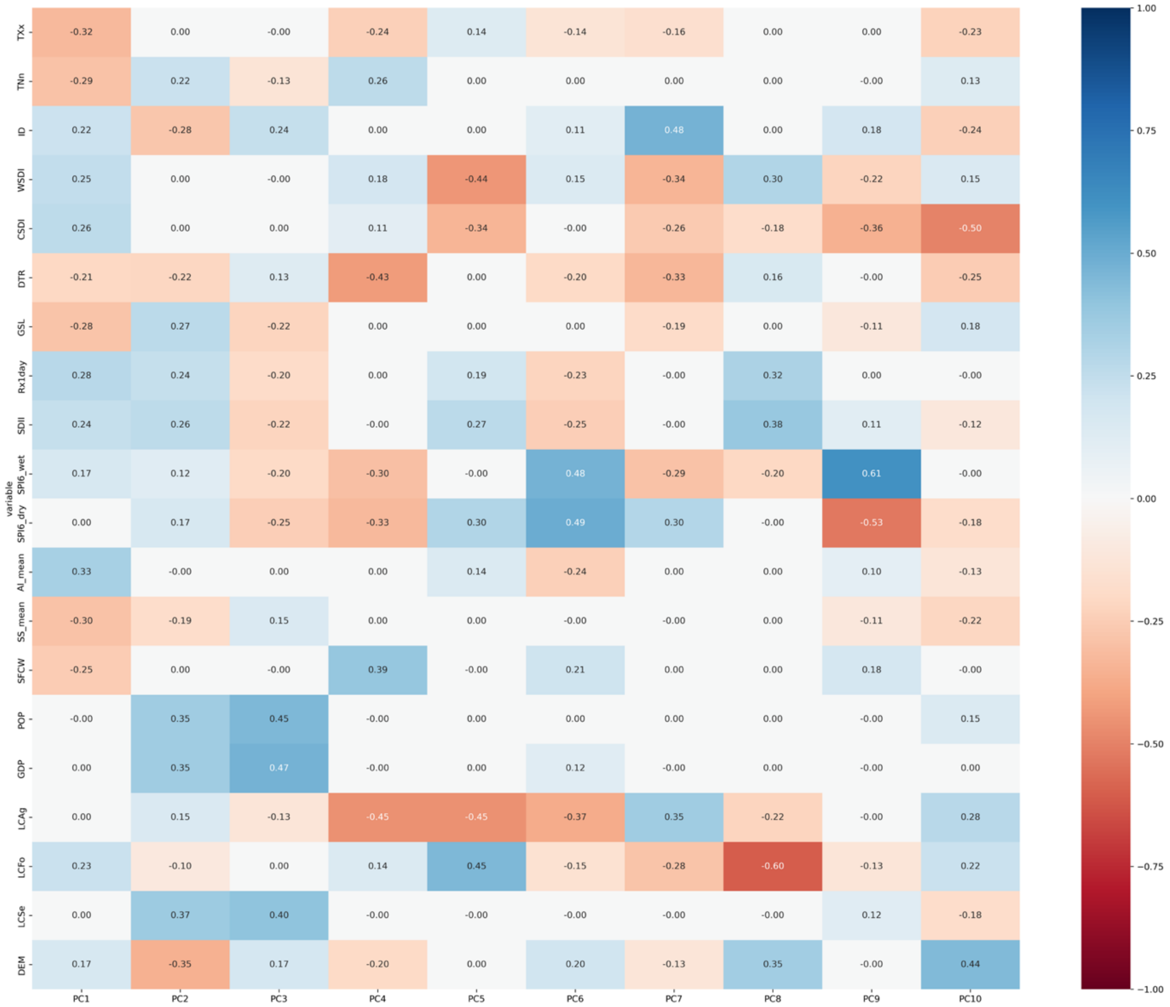


Figure 29 Outcome of the PCA analysis (heatmap) setting a threshold of 0.90 (10 PCs) to eventually consider to further reduce the original number of indicators.

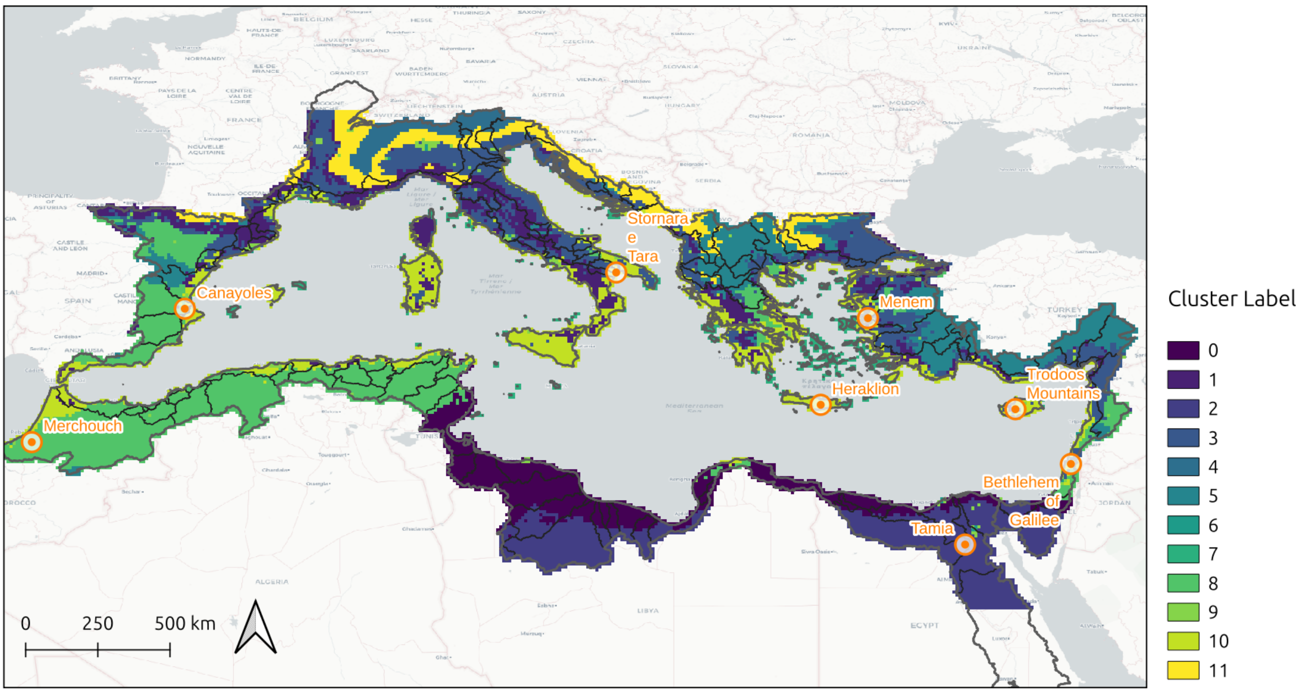


Figure 30 Outcome of the clustering analysis with k-means (k=12), which shows the map of the Mediterranean region divided into 12 clusters.

Several configurations have been tested and the parameters domain of the different runs is reported in the Annex I (Table A1), along with some correspondent outputs (e.g., input matrix composed not only with means but also trend values). As observable in Table A1, the ML procedure is set to successfully run also with projected future time series in the input matrix, useful to provide useful insights for orienting policy recommendations and support decision-making processes. Nevertheless, these aspects are explored in Task 6.3 and will be reported in D6.3 *LanDS scenarios to inform policy making*, in collaboration with TUC.

To support the interpretation of the clustering outputs, we have produced different boxplots looking at the cluster distribution (x-axis) against the single indicators (y-axis). E.g., population, GDP, and land cover correspondent to settlements are very high in cluster 7, while temperatures are very high (e.g., TXx, WSDI) while precipitations very low (PrAvg, CDD) in cluster 2, indicating more arid areas. Figure 31 and Figure 32 report some examples of the boxplots useful for the outputs' interpretation, while the boxplots correspondent to all ML tool inputs indicators of this application case are reported in Annex II.

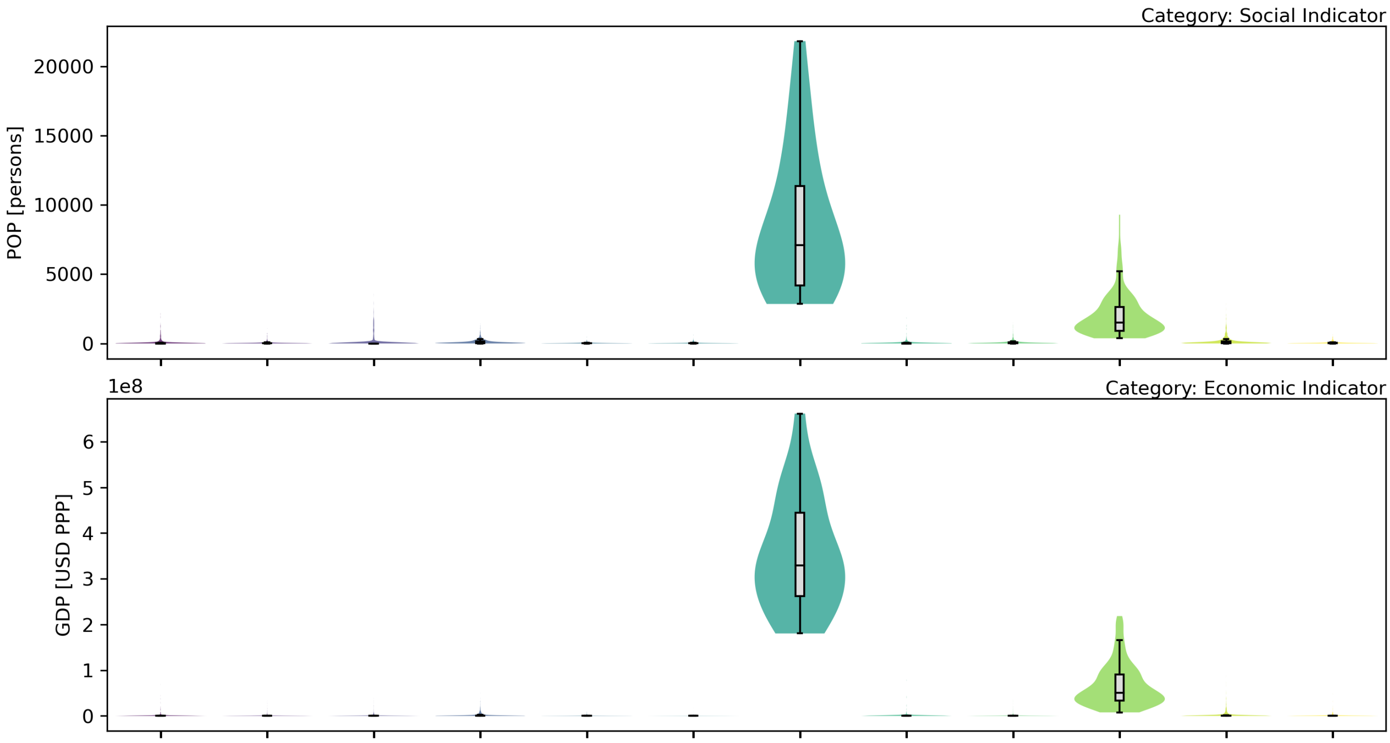


Figure 31 Boxplot of the clusters distribution (k=12 on the x-axis) for Population and GDP (y axis).

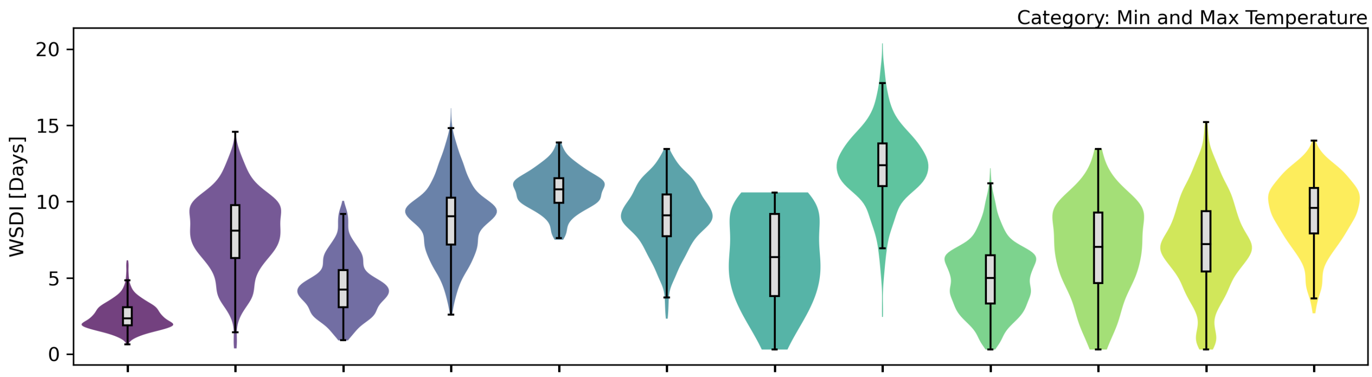


Figure 32 Boxplot of the clusters distribution (k=12 on the x-axis) for WSDI (y axis).

2.4.4 Expert-based filter

The outcomes of the clustering analysis (step 3 of the methodology of the ML procedure) identifying similar macro-areas in the Mediterranean region can be used as input for the expert-based filter analysis, which integrates higher resolution datasets and specific local conditions to find potentially suitable areas for each specific restoration action run in the PAs.

The expert-based filter is yet in development and will process the following steps:

- Load and preprocess higher resolution datasets
- Define expert-based criteria, collaborating with experts (PA leaders) defining thresholds and/or specific conditions based on the additional data (e.g., WOCAT factsheets)
- Apply the filter and plot results, scale/interpolate if needed
- Identify areas potentially suitable for out- and upscaling of restoration actions
- Discuss outcomes with local experts and eventually re-tune filter

Figure 33 reports a visual example of a possible implementation of this filter, based on filters and threshold not discussed with ERLL experts:

- areas identified as similar to PA2 (Heraklion, Greece) by the ML Tool (belonging to cluster no. 10 for the application case described in subsection 2.4.3);
- areas with given values of soil depth;
- areas within a certain range of elevation.

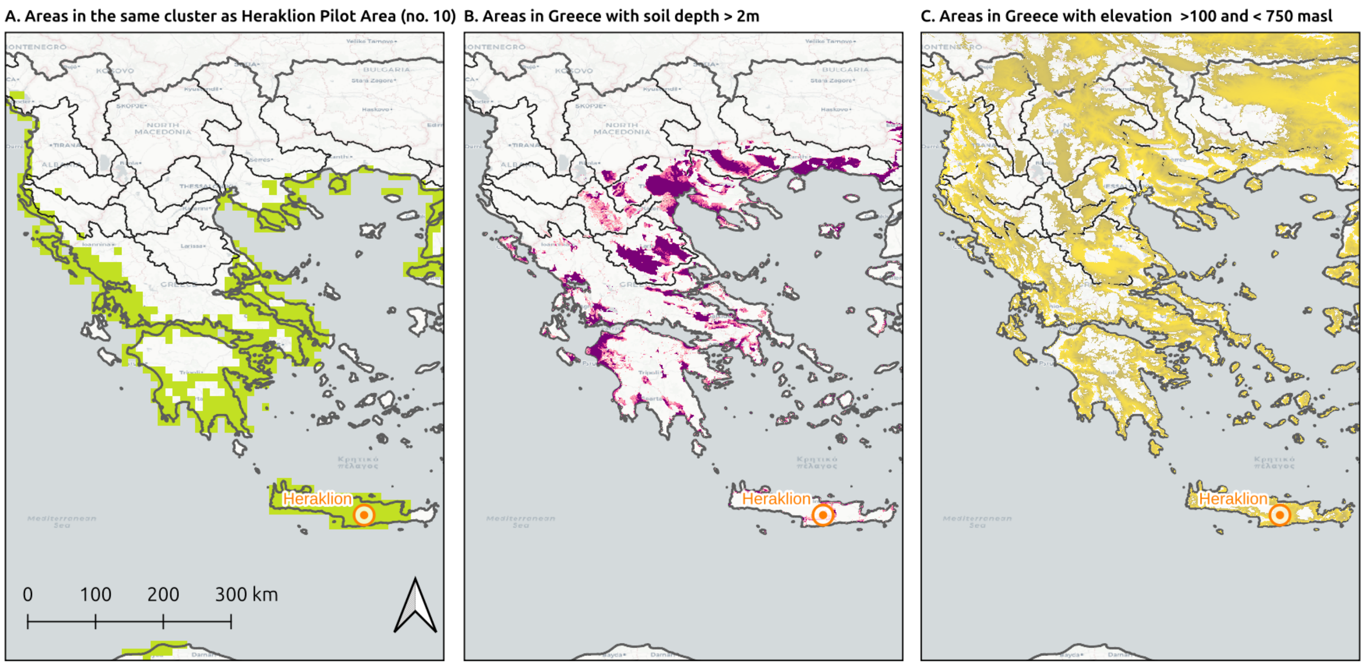


Figure 33 Example of expert-filter implementation for PA2 (Heraklion, Greece) with a zoom in Greece: map of areas belonging to cluster 10 which includes PA2 (left), map of areas with soil depth > 2m (centre); areas elevation higher than 100 and lower than 750 m asl.

The intersection of these three maps would identify a subset of areas to be investigated from ERLL experts or other stakeholders as potentially suitable for the Restoration Action proposed. The implementation of the expert-based filters and related outcomes will be reported in deliverable D6.3 *LanDS scenarios to inform policy making*.

3 Online documentation

The LanDS documentation[24] is available as an online page on the LanDS web interface, accessible by clicking on the "i" button on the top bar menu next to "Login" (see Figure 34).

In its current version, it includes a general description of the toolbox, the list of WP4 submitted deliverables, a brief presentation of the four available tools (repository, data viewer, indicators library, and ML tool) with correspondent links and guidance on the main objectives and functionalities. It is enriched by examples and screencasts to guide the user in specific actions (e.g., login, upload/edit data, select visualization types for different file formats).

This page will be periodically updated with refinements and additional information, if and when needed.

On December 13th, we run a workshop with selected LanDS users (one per PA), to guide them to the LanDS functionalities in general and the usage of its data monitoring system.


LanDS | Repository | Data viewer | Indicators | ML Tool | Dashboard | Log in

Documentation

The LanDS is a science-based Land degradation Decision-Support Toolbox targeted to support REACT4MED participants (partners, stakeholders, and policy makers), by providing a safe and effective georeferenced repository to store, share and reuse data collected by the pilot areas, and by implementing indicators to assess land restoration measures impacts across different geographic areas. Combining existing knowledge from public repositories and expertise coming from the pilot areas, the LanDS elaborates machine-learning (ML)-based procedures to identify critical areas in the Mediterranean, where to focus up or out-scaling of restoration measures, exploring future climate and socio-economic scenarios. Finally, the LanDS web dashboard, specifically addressed to decision makers, is in development to effectively share project's outcomes and support assessment and policy recommendation activities.

The LanDS is then composed of five different tools: 1) a geo-referenced data repository serving as a knowledge base by collecting site-specific data and resources from the pilot areas as well as from global or regional public repositories for the Mediterranean area; (2) a data viewer, containing a set of visual analytics tools linked with this repository allowing the effective sharing and access to data, and their interactive exploration; (3) an indicators library implemented as a modular and generalised code library applicable to different geographical contexts based on collected data and indicators identified; (4) a ML tool to identify potentially suitable areas in the Mediterranean for up- and outscaling of restoration measures; (5) a web interactive dashboard, providing a harmonised land degradation assessment and evaluation of impacts of restoration measures, as well as a user-friendly interface.

This online documentation servers as a user's guide to navigate the toolbox and getting started with its functionalities.

Before exploring the tools in more detail, please find the link to the submitted deliverables explaining the development process of the toolbox and correspondent functionalities:

- D4.1 LanDS requirements and specifications (Feb. 2023): Report summarising the outcomes of ERLL stakeholders' involvement process and defining features and requirements of the LanDS toolbox.
- D4.2 LanDS toolbox prototype (Apr. 2024): Prototype of LanDS toolbox, including first versions of the geo-referenced data repository, the visualization tools and the indicators library, ready to be shared with partners and stakeholders for testing and feedback gathering.
- D4.3 Final LanDS toolbox (Dec. 2024): Release of the final version of the LanDS toolbox, including the planned tools, populated with project data and documented by a user-guide with examples and tutorials. The software developed will be released in an online public repository.

(1) The georeferenced data repository can be access from the LanDS landing page and/or from the top bar menu, selecting:

- Map (LanDS landing page), which shows the map of the Mediterranean Basin with nine location pins to highlight the nine available data collections (including the 8 pilot areas in Cyprus, Greece, Italy, Spain, Morocco, Israel, Turkey, and Egypt, plus a general one for the Mediterranean Area) and, on the right-hand side, a column with the collections represented by their reference image, their title/location, and associated tag (issues and/or restoration action), which can be easily scrolled down and selected as needed: the user has the option to select the collection /from the pins on the map or from the reference images;
- Collections, displaying the 9 collections as described in the previous point;
- Data Browser, a dynamic table where datasets (resources and indicators) are grouped by collection and linked to their descriptive page (by clicking on the dataset name) and to the related dataset (by clicking on file's link). Other available functionalities are listed below:
  - Show (on the top-left side of the page): it displays the selected number of available entries;
  - Search (on the top-right side of the page): it provides a quick filter on displayed rows based on text included in each field;
  - Advanced Search (next to Search): performs a more refined spatio-temporal query across all the collections.

Check out this quick overview to first navigate the LanDS from the landing page to the repository and datasets.

Full open access, data visualization and download is granted to anyone visiting the LanDS website, while the possibility to upload and edit data is given only to logged in users (here is how to quickly login, if your account is activated by the admin).

(2) The Data Viewer provides quick and effective visualizations of the different datasets uploaded in the repository and can be accessed from:

- from the dataset/item page, to contextually show the visualizations available for the specific dataset;
- from the top menu bar, to directly open the Data Viewer page.

Figure 34 Screenshot of the first part of the LanDS page documentation already online.

[24] https://lands.soft-water.it/documentation

4 Technologies

Starting from expertise and tools developed in previous international projects by SoftWater members (e.g., DAFNE H2020 project, H2020 Project Ô), the LanDS toolbox has been developed relying on well-known and mature open-source software projects, schematised in Figure 35:

- Postgresql[25] + PostGIS[26] for the creation of relational GeoDatabase: one of the most advanced and modern relational database managers. Its spatial extension PostGIS allows the managing and processing of geographical data, while its capability to manage also JSON structured data allows an extra-flexibility for application development and data storage;
- STAC SpatioTemporal Asset Catalogs[27], a common language to describe geospatial information, so it can more easily be worked with, indexed, and discovered;
- Leaflet[28] for creating interactive web maps, exposing directly geographical data or complex map providedby map servers;
- FastAPI[29] as a web framework for building APIs with Python programming language, with a modern, high-performance and scalable approach;
- Python code libraries for indicators computation, geographical data management, interactive charts generation, machine-learning: pandas[30], geopandas[31], rasterio[32], xarray[33], plotly[34], hvplot[35], sci-kit learn[36], keras[37], dask[38];
- Drupal[39] as Content Management System: a framework that can be tailored and customized to create simple websites or complex web applications. It has several standard features out-of-the-box, like easy content authoring, reliable performance, and excellent security, but it gains popularity because of its flexibility and modules library, to expand functionalities and customize the appearance;
- Vue[40]: a JavaScript framework for building user interfaces customising Drupal aspect.
- Jupyter[41]: interactive development environment for notebooks, code, and data. Its flexible interface allows users to configure and arrange workflows in data science, scientific computing and machine learning;
- Sci-kit Learn[36]: a ML library for Python, providing simple and efficient tools for predictive data analysis.

25 https://www.postgresql.org/
26 https://postgis.net/
27 https://stacspec.org
28 https://leafletjs.com/
29 https://fastapi.tiangolo.com/
30 https://pandas.pydata.org/pandas-docs/stable/index.html
31 https://geopandas.org/en/stable/
32 https://rasterio.readthedocs.io/en/stable/
33 https://docs.xarray.dev/en/stable/
34 https://plotly.com/
35 https://hvplot.holoviz.org/
36 https://scikit-learn.org/stable/
37 https://keras.io/
38 https://www.dask.org/
39 https://www.drupal.org/
40 https://vuejs.org
41 https://jupyter.org/

The full stack of these components has also been open sourced as docker containers on the LanDS code repository[42].

Figure 35 Overview of technologies used for each one of the LanDS toolbox components.

42 https://gitlab.com/lands-r4m/lands_web

5 Conclusions

This deliverable describes the final version of WP4 scientific Land degradation Decision-Support Toolbox, which contributes to the REACT4MED scope of up- and outscaling the potential application of the land restoration actions promoted in the Ecosystem Restoration Living Labs to the Mediterranean area. The LanDS aims to be a comprehensive innovative toolbox integrating data storage, computational, and visualisation tools, with a ML-based procedure for the identification of target restoration areas.

As demonstrated in this report, the LanDS combines existing knowledge from public repositories and expertise from the ERLLs in the pilot areas, supporting better-informed land restoration actions and implementing a ML procedure to identify suitable areas for their potential up- and outscaling in the Mediterranean region. In the upcoming period, our target is also to explore future climate and socio-economic scenarios, to elaborate policy recommendations (together with WP6) and to identify investment opportunities for public and private actors, based on the criteria of maximum cost-effectiveness and impact (WP5). Finally, open access to the decision support tools developed will support restoration actions' upscaling beyond the lifetime of the project.

In a couple of months (M34), we will complete D4.4 *LanDS Dashboard* and work in collaboration with TUC on elaborating D6.3 *LanDS scenarios to inform policy-making*, due in M36, in which we will include the work on the expert-based filters and future scenarios. A final webinar on the LanDS dashboard (organised jointly with WP7) will be held, to present the final LanDS toolbox to all stakeholders and policy-makers involved in the ERLLs and launch the LanDS dashboard.

References

Allam, A. et al., 2020: Specific climate classification for Mediterranean hydrology and future evolution under Med-CORDEX regional climate model scenarios. Hydrology and Earth System Sciences 24(9), 4503–4521. doi:10.5194/HESS-24-4503-2020.

Calka B. et al. 2022: The Ratio of the Land Consumption Rate to the Population Growth Rate: A Framework for the Achievement of the Spatiotemporal Pattern in Poland and Lithuania. Remote Sensing 14(5):1074, doi: 10.3390/rs14051074.

Forgy, E. W., 1965: Cluster analysis of multivariate data: efficiency versus interpretability of classifications, Biometrics, 21, 768–769.

IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2391 pp. doi:10.1017/9781009157896.

Jolliffe, I. T. (2002). Principal Component Analysis. Springer Series in Statistics. Springer-Verlag, New York.

Lehner, B., Grill G. (2013). Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. Hydrological Processes, 27(15): 2171–2186, doi: 10.1002/hyp.9740.

MacQueen, J., 1967: Some methods for classification and analysis of multivariate observations, in: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, Univ. of Calif. Press.

Wang. & Sun, F. (2023). Global gridded GDP under the historical and future scenarios [Dataset], Version 7. In *Zenodo*.

Yang, Y. et al. (2019). Hydrologic implications of vegetation response to elevated CO2 in climate projections. Nature Climate Change 9(1), 44–48, doi: 10.1038/s41558-018-0361-0.

Yang, D. et al. (2023). Projecting spatial interactions between global population and land use changes in the 21st century. Npj Urban Sustainability, 3(1), doi: 10.1038/s42949-023-00131-y.

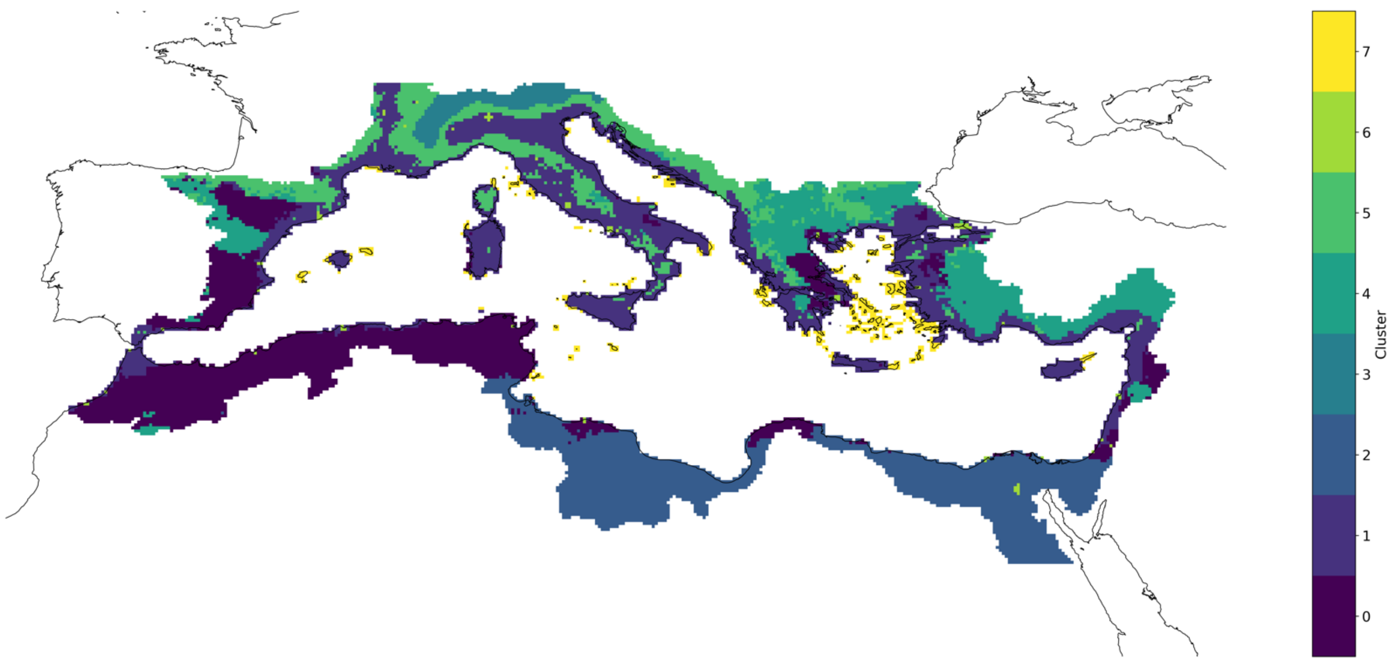
Annex I: ML tool runs

Table A1 reports the parameters domain of all ML tool tests, which have been run in the past months:

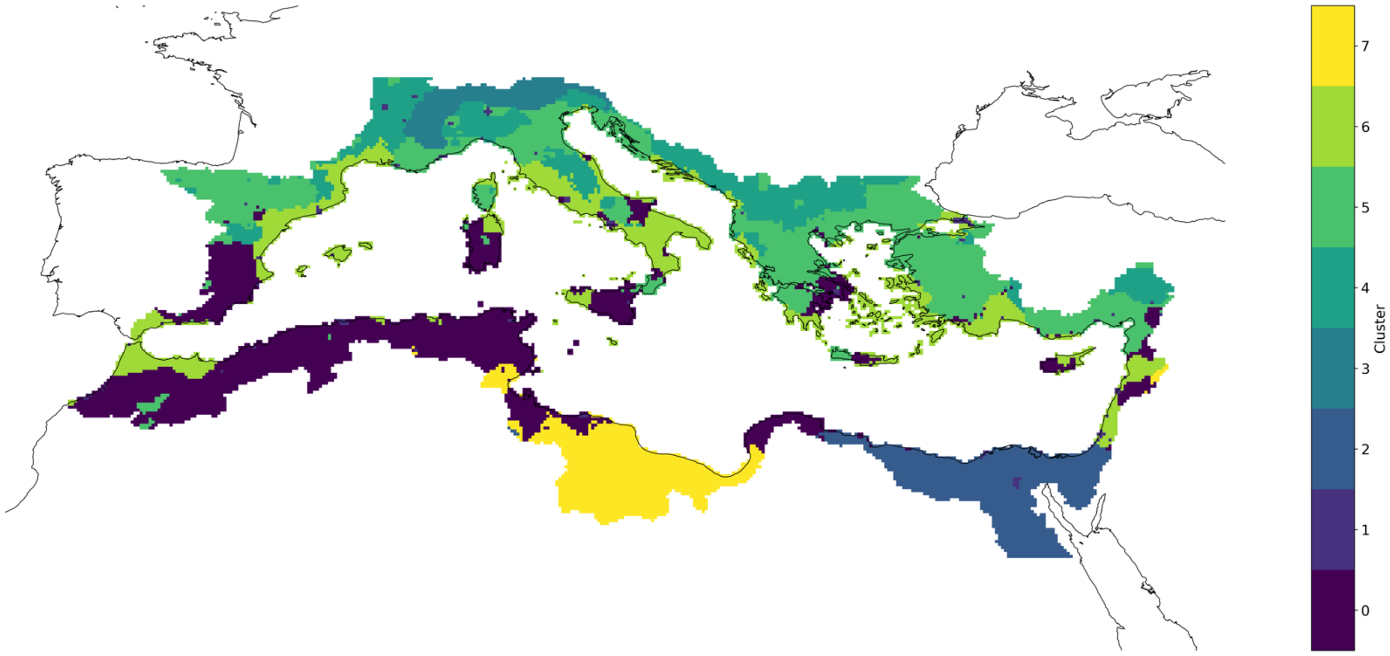
Input matrix configuration: time (rows)	• 2001-2019 (past), aggregated over the full time horizon or every five years • 2025-2100 (future), aggregated over the full time horizon or every five years
Input matrix configuration: indicators values (columns)	• Mean • Mean + Trend • Mean + Trend + Variance
Correlation analysis	50% - 90%
Correlation analysis with categories binding (see Table 4)	Yes/No
#Input indicators	15 - 56 (based on different levels of selection)
#Clusters outputs	k = 2 - 20
Future scenarios available	SSP245, SSP585

The following figures report additional tests run:

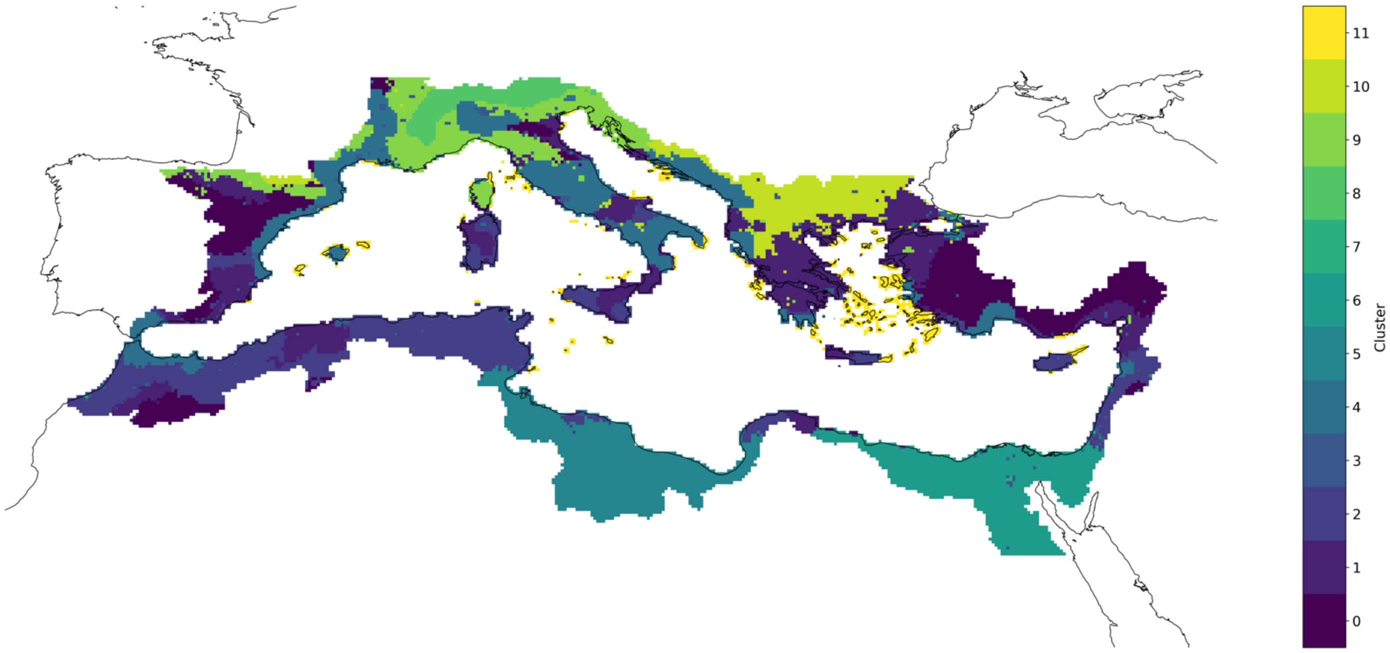
- Application as in 2.4.3, but with k=8



- Application as in 2.4.3, but considering mean + trend in as input indicators and k=8



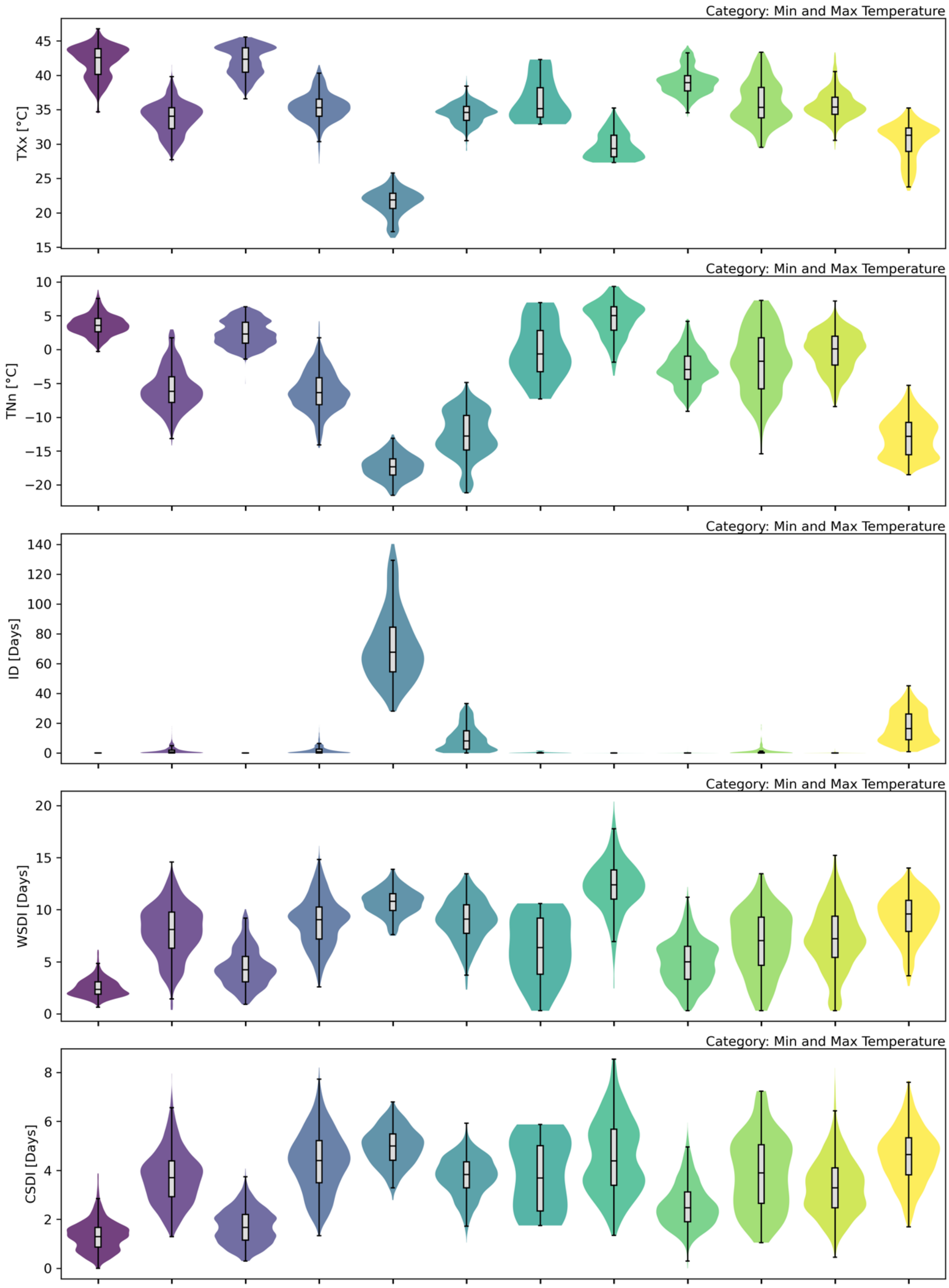
- Application as in 2.4.3, but considering mean + trend in as input indicators and k=12



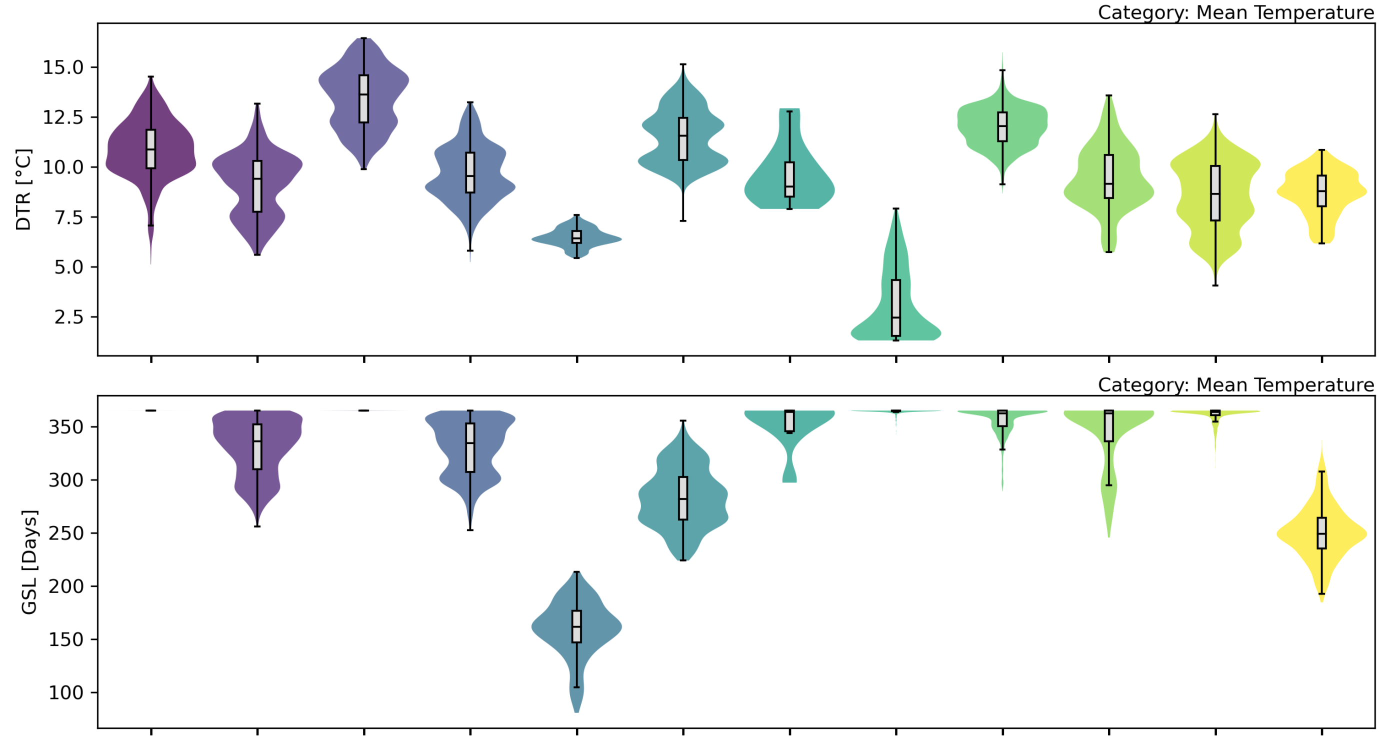
Annex II: ML tool clusters distributions on input indicators

The ML tool clusters distribution against their input indicators, resulting from the application case of subsection 2.4.3 (k=12, x-axis), are reported in the following:

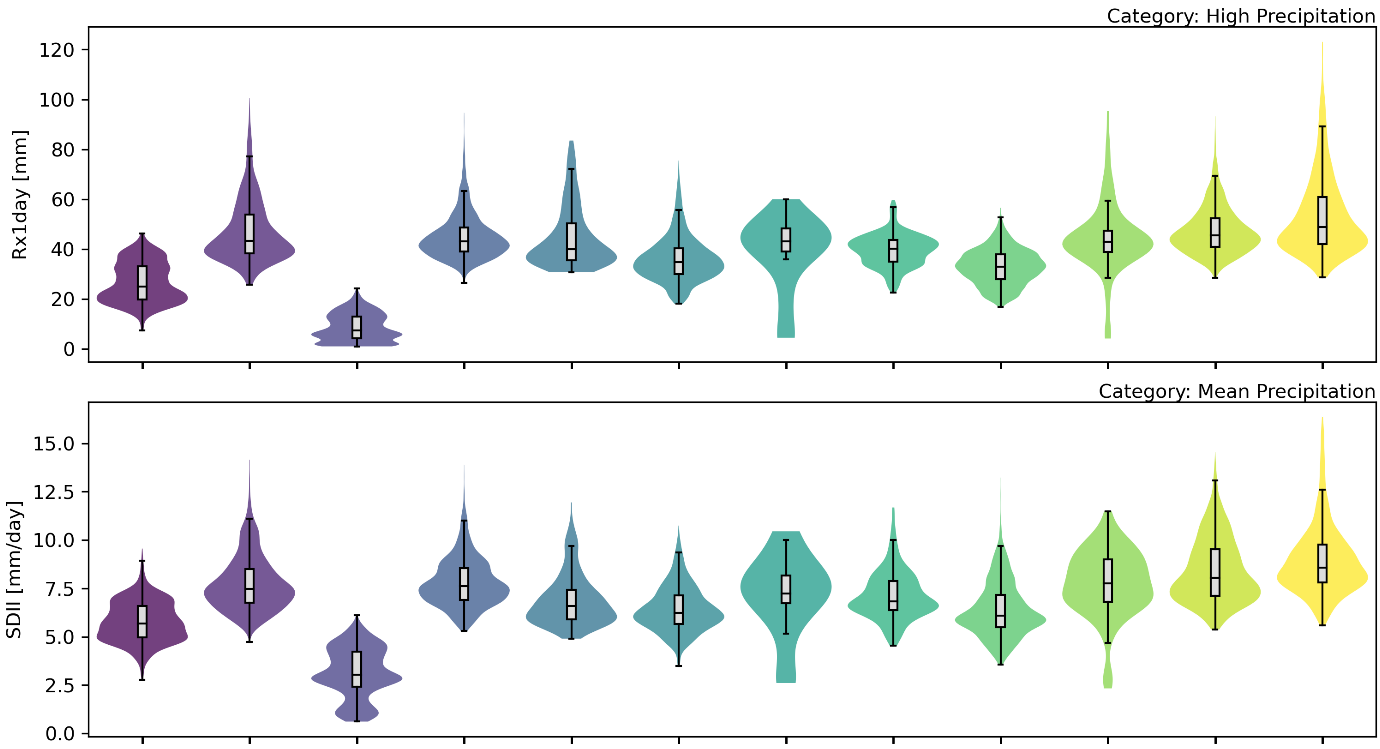
- Category: Mean and Max Temperature



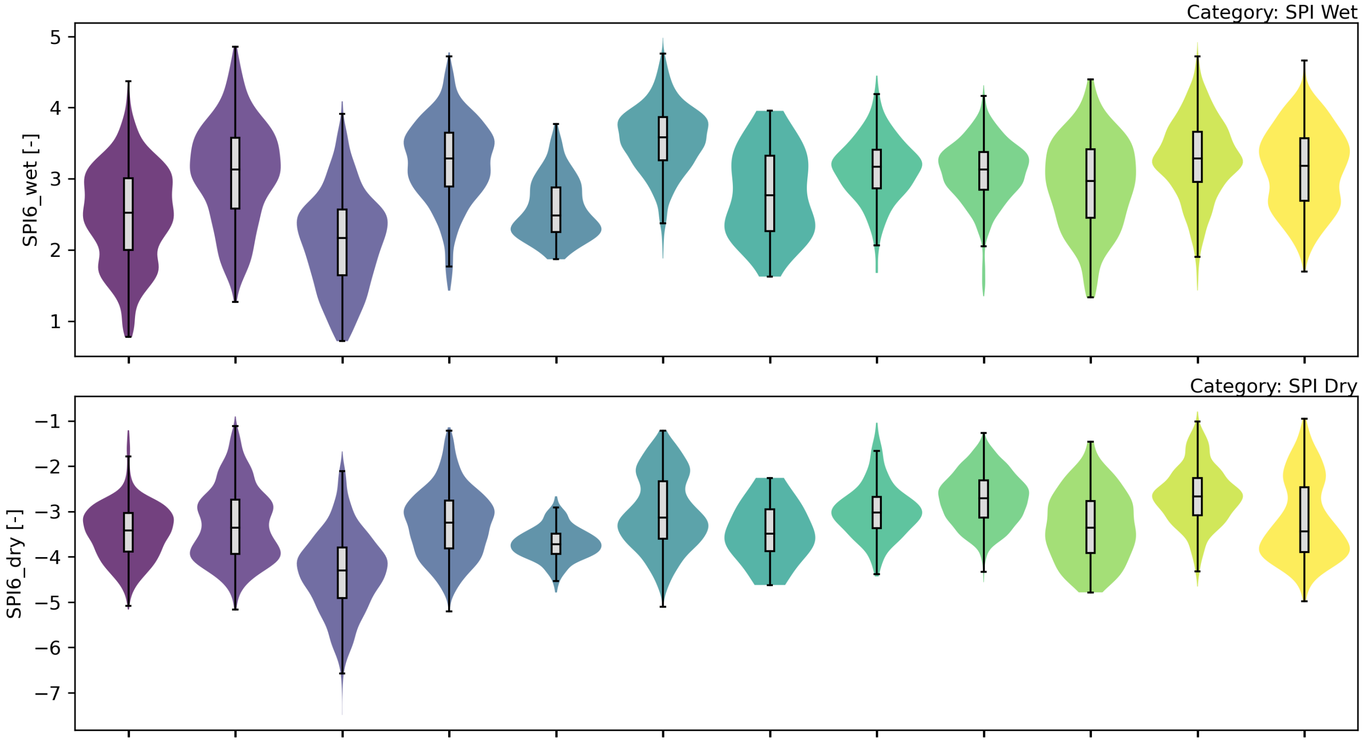
- Category: Mean Temperature



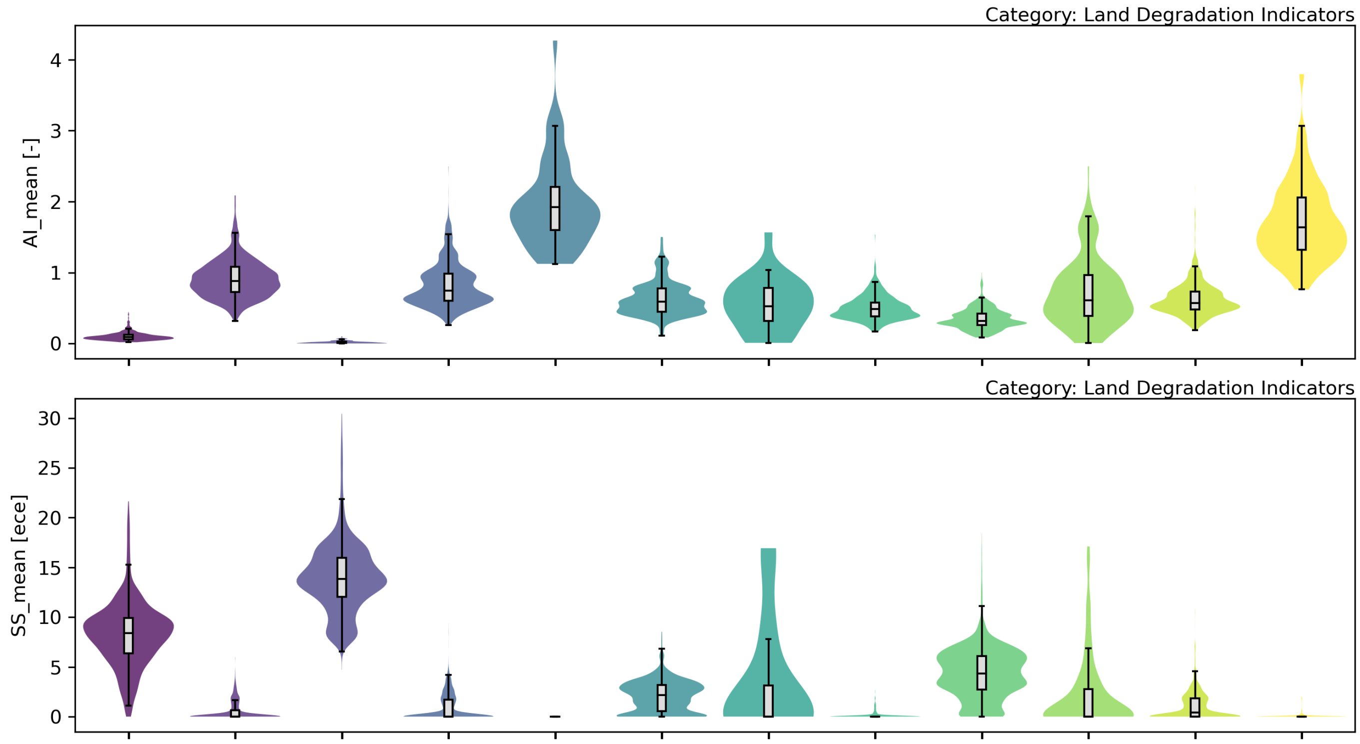
- Categories: High Precipitation and Mean Precipitation



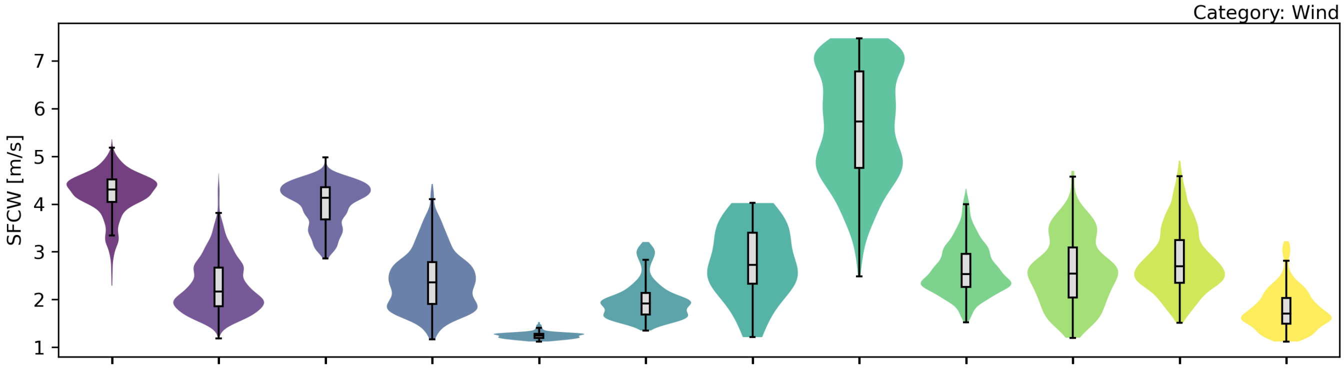
- Categories: SPI Wet and Dry



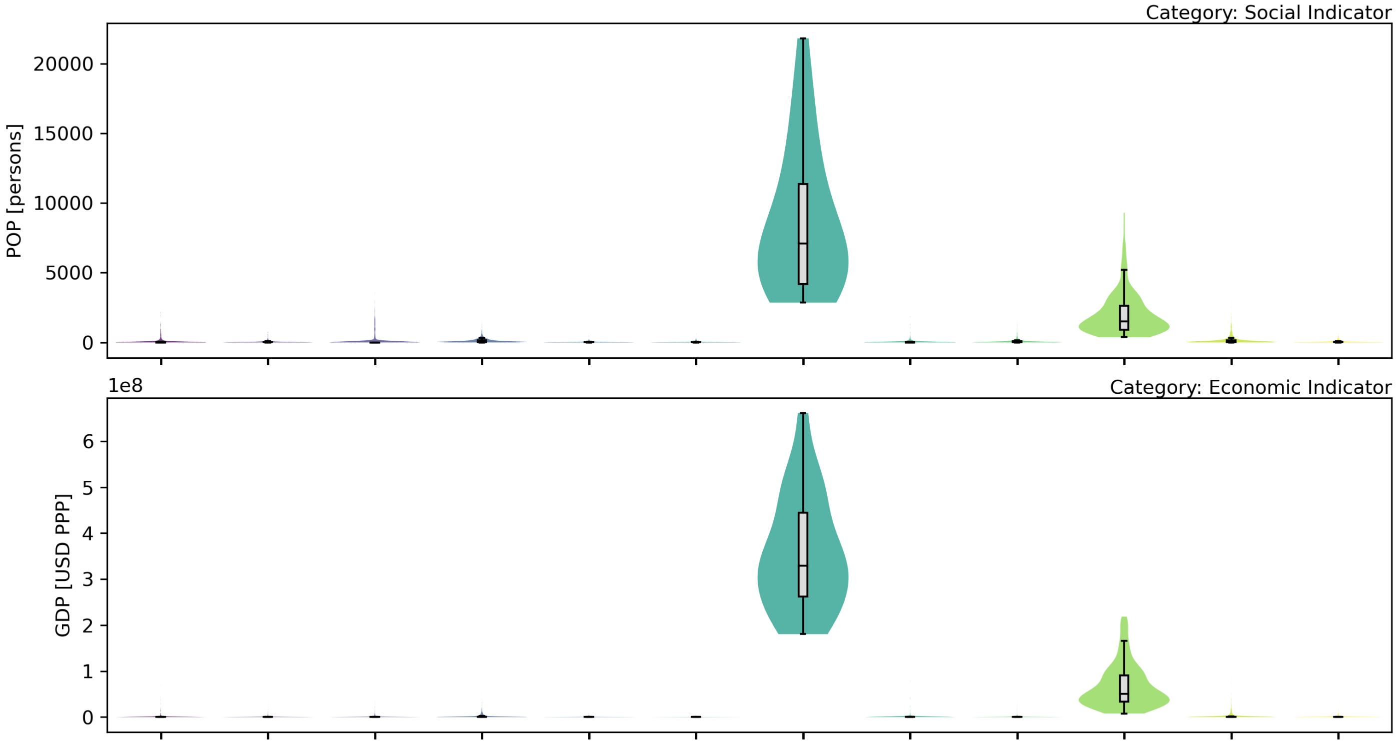
- Category: Land degradation indicators



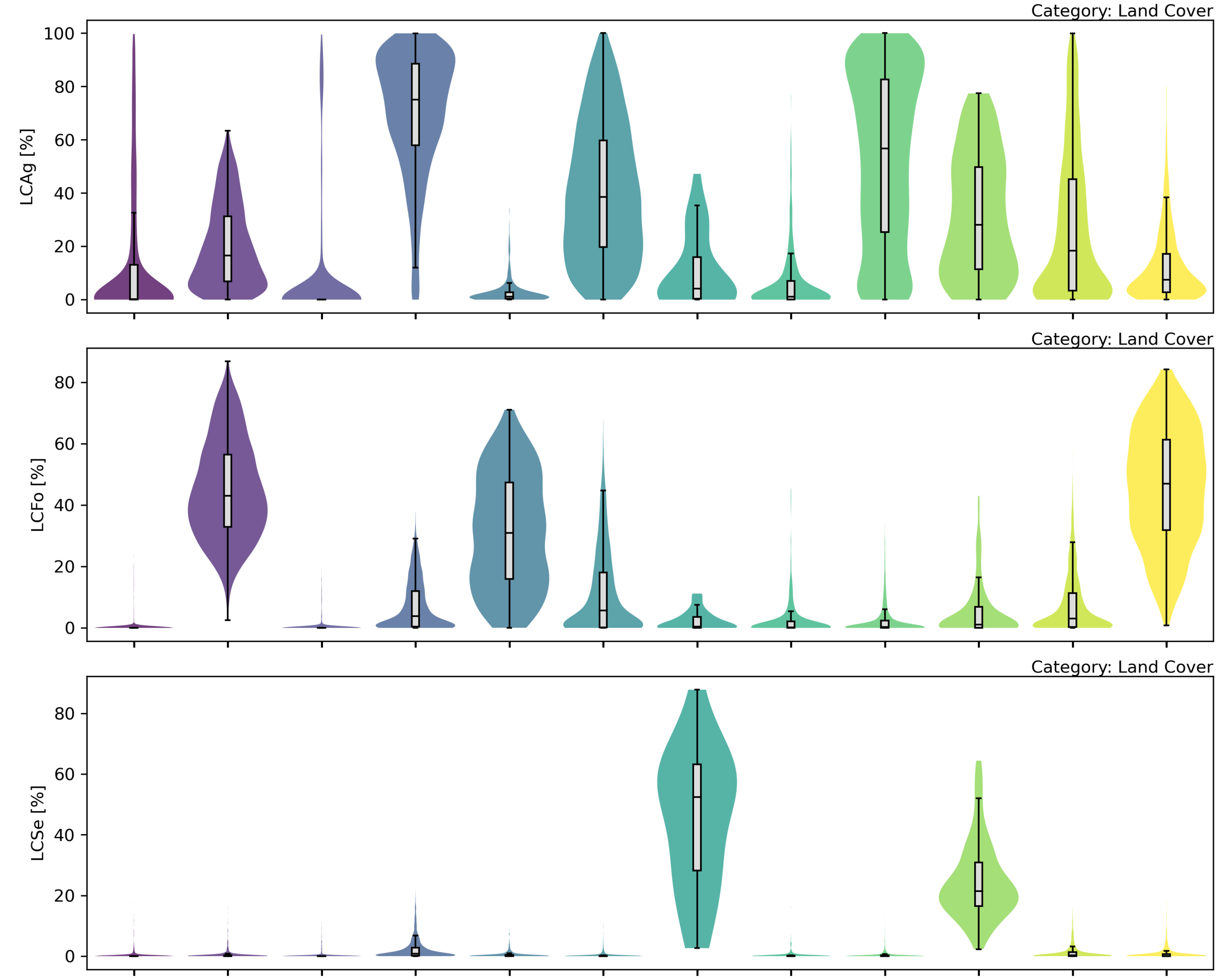
- Category: Wind



- Categories: Social and Economic Indicators



- Category: Land Covers indicators



- Category: Topography indicators