



Cyprus Earth Observation Data Cube towards water resource management in the context of climate crisis

Stelios P. Neophytides, Gregory Giuliani, Charalampos Kontoes, Diofantos G. Hadjimitsis & Michalis Mavrovouniotis

To cite this article: Stelios P. Neophytides, Gregory Giuliani, Charalampos Kontoes, Diofantos G. Hadjimitsis & Michalis Mavrovouniotis (21 Mar 2026): Cyprus Earth Observation Data Cube towards water resource management in the context of climate crisis, Big Earth Data, DOI: [10.1080/20964471.2026.2646028](https://doi.org/10.1080/20964471.2026.2646028)

To link to this article: <https://doi.org/10.1080/20964471.2026.2646028>



© 2026 The Author(s). Published by Taylor & Francis Group and Science Press on behalf of the International Society for Digital Earth, supported by the International Research Center of Big Data for Sustainable Development Goals.



Published online: 21 Mar 2026.



Submit your article to this journal [↗](#)



Article views: 780



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

RESEARCH ARTICLE

 OPEN ACCESS



Cyprus Earth Observation Data Cube towards water resource management in the context of climate crisis

Stelios P. Neophytides^{a,b}, Gregory Giuliani^{d,e}, Charalampos Kontoes^f,
Diofantos G. Hadjimitsis^{b,c} and Michalis Mavrovouniotis^{a,b}

^aBig Earth Data Analytics, ERATOSTHENES Centre of Excellence, Limassol, Cyprus; ^bDepartment of Civil Engineering and Geomatics, Cyprus University of Technology, Limassol, Cyprus; ^cERATOSTHENES Centre of Excellence, Limassol, Cyprus; ^dInstitute for Environmental Sciences, EnviroSPACE Lab, University of Geneva, Geneva, Switzerland; ^eInstitute for Environmental Sciences, GRID-Geneva, University of Geneva, Geneva, Switzerland; ^fBeyond Centre of EO Research & Satellite Remote Sensing, National Observatory of Athens, Athens, Greece

ABSTRACT

The amount of Earth observation data is getting larger day by day. This rapid evolution of the field and the exponential growth of data generation mediums urges the need for modern ways of fast analysis and processing of satellite images. Earth Observation Data Cubes enhance the way Earth data are distributed, handled, stored, and analyzed by providers and users. In this study the development of the National Earth Observation Data Cube of Cyprus is described and the importance of such software-based infrastructure is demonstrated through a water resource monitoring use case in the context of climate crisis in the Mediterranean region. Moreover, this study exhibits the capabilities of semantic classification in Earth observation, introducing a spatially generalized approach for easier environmental monitoring. Such developments can enhance the preservation of the environment and its protection from future natural and human disasters.

ARTICLE HISTORY

Received 15 December 2025
Accepted 7 February 2026

KEYWORDS

Remote Sensing; Earth Observation; big data; data cube; climate crisis

1. Introduction

Earth observation (EO) allows us to monitor changes and processes on Earth's surface. By properly analyzing the large volumes of EO data, we can gain valuable insights into different aspects of our planet. More specifically, the continuous data acquisition from satellites and other remote sensing (RS) platforms creates opportunities for systematic monitoring of agriculture, land, water, and atmospheric dynamics among others, at different spatial and temporal resolutions (Ajibola & Cabral, 2024; Allu & Mesapam, 2025; Sun et al., 2024). The rapid expansion of EO missions has, however, led to an exponential growth of data volumes and variety. For instance, NASA's Landsat Archive consists of more than 10.2 million images

CONTACT Stelios P. Neophytides  stelios.neophytides@eratosthenes.org.cy  ERATOSTHENES Centre of Excellence, 82 Franklinou Roosevelt, Limassol 3012, Cyprus

© 2026 The Author(s). Published by Taylor & Francis Group and Science Press on behalf of the International Society for Digital Earth, supported by the International Research Center of Big Data for Sustainable Development Goals. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

exceeding 9 petabytes of data while the European Space Agency's Sentinel mission has already produced much more data. An amount of more than 455 million satellite datasets have been downloaded globally in 2022 alone, underscoring the accelerating demand for EO information.¹ In general, 2024 studies showed that current EO datasets reached a total of 807 petabytes and more than 100 petabytes are generated every year (Wilkinson et al., 2024).

Thus, the recent uptake of EO data is raising barriers for efficient storage, processing, and analysis. Conventional approaches are insufficient for timely and actionable information extraction from such vast archives. To address this, Earth Observation Data Cubes (EODCs) have emerged as a strong paradigm, offering standardized, analysis-ready structures enabling providers and users with coherent handling, storage, and usage of EO data. Although the EODCs' concept is not novel, their national development and support are highlighted (Dhu et al., 2019).

Moreover, recent advances in EODCs have moved beyond simple pixel-based processing and analysis to incorporate semantic frameworks, which assign meaningful land cover or land use classes to each pixel based on spectral, spatial, and temporal information (Kröber et al., 2025). Semantic querying and classification enables the translation of raw satellite data into interpretable concepts. In addition, integrating semantics, EODCs can support automated large-scale assessments of environmental changes and phenomena, improve the handling of ambiguous pixels, and enhance interoperability for downstream applications.

Countries under the threat of climate crisis, like Cyprus which is located in the eastern Mediterranean region, stand to benefit significantly from EODCs and semantic-based approaches. The island faces recurrent droughts and water shortages making systematic monitoring of water resource management a fundamental need (Montero et al., 2024; Neophytides et al., 2024). A national EODC can serve as a software-based infrastructure to strengthen environmental resilience, support evidence-based policies, and mitigate risks from both natural and human-induced hazards (Eliades et al., 2023).

This study contributes to the broader challenge of water security in the Mediterranean region, where hydrological monitoring is limited, but decision support needs are high. CEODC aims to enable automated monitoring of reservoir dynamics, supporting risk assessment in water resources.

In this study, we present the development of a national EODC for Cyprus (CEODC) and demonstrate its capabilities through water resource monitoring case study in the context of the climate crisis. The contributions of this study are: the i) design and development of the CEODC, ii) a semantic-based framework for water resource monitoring, and iii) analysis of CEODC's interoperability.

The remainder of the paper is structured as follows. [Section 2](#) provides an overview of previous works related to Earth Observation Data Cubes and their utilization for water resource management and monitoring, [Section 3](#) describes the design and development of CEODC and provides an overview of indexed satellite products. Moreover, [Section 4](#) outlines study area of the demonstration use case, ground-truth datasets, semantic framework, and methodologies used. [Section 5](#) presents the results of the semantic-based water resource monitoring framework and the CEODC's user interfaces. Finally, [Section 6](#) discusses findings and CEODC's interoperability based on six faces and [Section 7](#) concludes this work.

2. Related works

EODCs have gained increased attention for managing large EO archives through multi-dimensional raster databases. Towards this direction, various frameworks for the development of EODCs have been proposed by the scientific community.

The first major approach for multidimensional database management for raster data was RasDaMan approach (Baumann et al., 1998). The framework was introduced as a streamlined store manager of images in terms of spatial and temporal resolution enabling faster analysis (Baumann et al., 1997). Another early framework, TerraLib, extended these capabilities with vector data handling alongside rasters (Câmara et al., 2008). Modern frameworks like xcube (xcube-dev, 2025) which is based on Zarr storage for multi-dimensional geospatial images (Khan et al., 2024), and Open Data Cube (ODC), which organizes metadata of EO products in a PostgreSQL/PostGIS relational database system (Killough, 2018) enhanced EO data management. One of the main attributes of EODC frameworks is their ability to serve users with Analysis Ready Data (ARD) (Killough, 2019; Killough et al., 2020).

In addition, frameworks like “Earth in a box” showed promising advancements towards light-weight EODCs adaptation for land surface and atmospheric variables (Gans et al., 2018). SpatioTemporal Asset Catalogs exposed EO data via catalogs ordered by spatial and temporal resolution (Hanson, 2019). Finally, ml4xcube demonstrated straightforward adaptation of EODCs into machine learning pipelines (Peters et al., 2025), while lexcube focused on immersive visualizations of EODCs (Söchting et al., 2024).

These frameworks represent different architectural strategies, ranging from array database engines like RasDaMan, to data management systems that handle both rasters and vectors like TerraLib, cloud-based chunked storage such as xcube, and metadata-driven relational systems like ODC.

Building on these core frameworks, national and regional EODCs have been developed worldwide. First, the Australian EODC served as pioneers and developers of ODC framework (Lewis et al., 2017). More countries and regions followed, such as Switzerland (Chatenoux et al., 2021), Brazil (Picoli et al., 2020), China (Cao et al., 2022), Armenia (Asmaryan et al., 2019), and Africa (Mubea et al., 2024) among others. In addition, Austria exhibits an ODC extension with semantic translation of pixels to land cover classes (Augustin et al., 2019).

EODCs demonstrated their applicability in different environmental fields such as agriculture (Drivas et al., 2022), snow coverage (Poussin et al., 2025), land cover (Chaves et al., 2021), floods (Wang & Gao, 2022), maritime (Kavouras et al., 2024), wildfire (Mantas et al., 2024), and extraction of essential climate variables (Augustin et al., 2021) among others. EODCs’ importance for understanding water security as a social-ecological challenge was previously highlighted (Bolognesi et al., 2018). However, applications in water resource monitoring remain comparatively limited. Major contributions include the Water Observation from Space (WOfS) product, first applied in Australia (Mueller et al., 2016) for multi-decadal water extent monitoring and later in Africa (Halabisky et al., 2024). WOfS uses a decision tree algorithm for water extent detection using Landsat with 93% accuracy. A SAR-based framework detects water bodies with 97% correlation to WOfS using machine learning on Sentinel-1 VV and VH polarizations (Kreiser et al., 2018). Optical and radar data stored in Digital Earth Africa were fused to overcome high cloud coverage

for water body detection in Ethiopia (Peppa et al., 2023). Finally, an alerting system for inland water algal blooms used Landsat data from an EODC (Malthus et al., 2019).

While these contributions demonstrate the value of EODCs in water resource monitoring, applications remain underrepresented, particularly in climate-threatened Mediterranean regions (Claro et al., 2024; Eliades et al., 2023). Our approach generalizes water resource monitoring through EODCs by using globally uniform Sentinel-2 SCL and a semantic-based methodology instead of traditional pixel-based classification. In addition, ODC was selected for CEODC due to its maturity, open-source model, and widely proven capabilities in national deployments.

3. Development of Cyprus Earth Observation Data Cube

Cyprus Earth Observation Data Cube (CEODC) hosts data for the entire island of Cyprus and a big part of Cyprus' Exclusive Economic Zone. Cyprus is an island in the eastern Mediterranean Sea (35 °N, 34 °E) and is part of the wider area of the Eastern Mediterranean, Middle East, and North Africa region (Eliades et al., 2023) as shown in Figure 1. The island covers an area of 9,254 square kilometers. To cover Cyprus, we need 5 Sentinel-2 tiles (i.e., T36SVE, T36SWE, T36SXE, T36SVD, and T36SWD) and the h20v05 MODIS tile. Unlike Sentinel-2 and MODIS, Sentinel-1 data are not distributed in fixed tiles but in orbit-based frames of approximately 250 kilometers swath width. Cyprus is intersected by two ascending and two descending relative orbits (i.e., 13, 36, 86, and 109), each providing overlapping coverage of the island. Table 1 briefly describes indexed satellite data by providing information of name, description of observations, space-borne instrument, spatial and temporal resolutions and year of launch. Figure 2 illustrates a visual abstract of the acquisition, pre-processing and indexing pipelines of CEODC.

Different acquisition pipelines are set-up for the data collection upon data availability from the different providers of each different satellite dataset. For the multiple NASA's

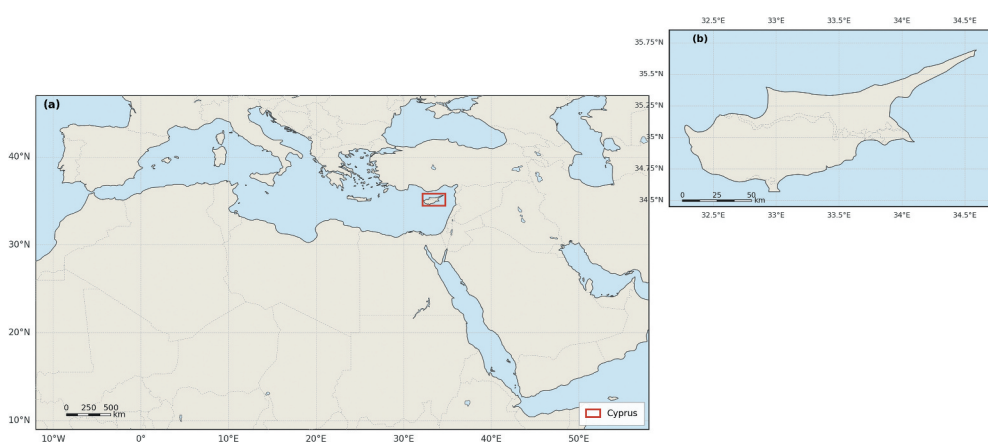


Figure 1. Cyprus' location in the wider area of eastern Mediterranean, Middle East, and North Africa region (a). Cyprus is pointed in the area with a red box. On the top right corner of the figure a zoomed map of Cyprus is shown (b). The x and y axis represents the coordinates. The figure was created with *cartopy* Python package.

Table 1. Earth observation products of Cyprus data cube. LAI stands for leaf area index, FPAR denotes fraction of photosynthetically active radiation, LST is land surface temperature, SAR stands for synthetic aperture radar and MSI for multispectral instrument.

Code Name	Description	Instrument	Spatial Resolution	Temporal Resolution	Launch
MOD16A2	Net Evapotranspiration	MODIS/Terra	500 m	8-Day	2021
MYD16A2	Net Evapotranspiration	MODIS/Aqua	500 m	8-Day	2021
MOD13Q1	Vegetation Indices	MODIS/Terra	250 m	16-day	2000
MYD13Q1	Vegetation Indices	MODIS/Aqua	250 m	16-day	2002
MOD44B	Vegetation Continuous Fields	MODIS/Terra	250 m	Annual	2000
MOD15A2H	LAI/FPAR	MODIS/Terra	500 m	8-day	2000
MYD15A2H	LAI/FPAR	MODIS/Aqua	500 m	8-day	2002
MOD17A2H	Gross Primary Productivity	MODIS/Terra	500 m	8-day	2000
MYD17A2H	Gross Primary Productivity	MODIS/Aqua	500 m	8-day	2002
MOD11A1	LST/Emissivity	MODIS/Terra	1 km	Daily	2000
MYD11A1	LST/Emissivity	MODIS/Aqua	1 km	Daily	2002
MOD14A1	Thermal Anomalies/Fire	MODIS/Terra	1 km	Daily	2000
MYD14A1	Thermal Anomalies/Fire	MODIS/Aqua	1 km	Daily	2002
Sentinel-1	Radar images	C-band SAR	5 m–20 m	12 days	2014
Sentinel-2	Optical images	MSI	10 m, 20 m, 60 m	10 days	2015

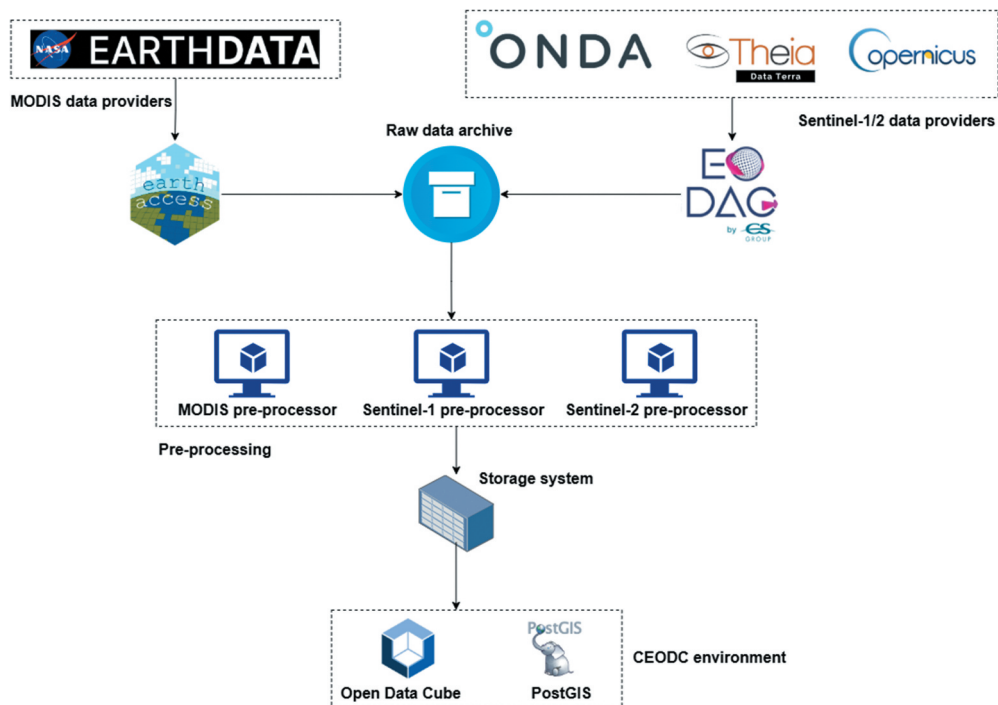


Figure 2. CEODC's acquisition, pre-processing, and indexing visual abstract.

MODIS products which are stored in CEODC, a connection to NASA's Earthdata via earth-access Python library is established (López et al., 2024). This specific library eases the way NASA's data can be accessed and is establishing an API connection for the data downloading. For the collection of ESA's Sentinel-1 and Sentinel-2 images, two different instances of Earth Observation Data Access Gateway (EODAG) are called (CNES and contributors, 2023). EODAG is responsible to serve users as a consolidated API through

Python scripting language. For each one of the Sentinel products, a fallback connection to three data providers (i.e., ONDA, Theia, Copernicus Data Space Ecosystem) is established. With this, redundancy regarding the successful download of each new image is achieved.

Each of the MODIS' products mentioned in Table 1 consists of different layers and each one contains different kinds of information related to the description. All the acquired products are in version 6.1. Moreover, some of the products' temporal resolution is denoted as 8-Day or 16-Day which has a different meaning for each dataset. For example, for MOD13Q1, each 16-day composite contains pixels with the highest quality based on cloud coverage, viewing geometry, and index value. For more information for each dataset, one can use the NASA's Earthdata Search Data Catalog.² Terra and Aqua satellites of NASA which both carry the MODIS instrument differ on the time they are crossing the equator. Terra is crossing equator during morning while Aqua crosses equator during afternoon. Both in tandem are covering Earth twice a day.

Sentinel-1 has a temporal resolution of 12 days but the constellation of 2 (i.e., satellite A and C) can decrease the temporal resolution to 6 days. On the other hand, Sentinel-2 revisits equator every 10 days and the constellation's temporal resolution is 5 days. Sentinel-1 images are capturing data in two different polarizations, the Vertical-Vertical (VV) and Vertical-Horizontal (VH) in a 10 meter resolution.

Data curation for all of the satellite data index in CEODC includes completeness and data quality checks, sensor (i.e., Sentinel-1, Sentinel-2, MODIS) specific data pre-processing, and conversion to Cloud Optimized GeoTIFFs (COGs) to generate Analysis Ready Data (ARD).

More specifically, MODIS' products are collected in a Hierarchical Data Format which is responsible for the management of complex scientific big data. Those are in a sinusoidal grid geographical projection and thus the fusion with other datasets is more complex. MODIS preprocessing consists of the direct conversion of each HDF variable to GeoTIFFs while ensuring no loss of data and metadata values.

Synthetic Aperture Radar images acquired by Sentinel-1 require a sequence of pre-processing steps before analysis, which consist of four main steps (i.e., radiometric calibration, speckle filtering, geometric correction, and conversion to decibel (dB) values). The whole pre-processing is done through snappy (European Space Agency, 2023) which is the Python API of ESA's Sentinel Application Platform software. Radiometric calibration was conducted using the auxiliary file and exporting the sigma band for the VV and VH polarizations, ensuring the ARD format of the Sentinel-1 acquisition. Speckle filtering was applied using the Lee filter with size 5×5 , a damping factor of 2 and an edge threshold of 5000 and is effectively decreasing the speckle noise. Geometric correction was carried out using the Shuttle Radar Topography Mission 1 arc sec height digital elevation model with bi-linear interpolation and a pixel spacing of 15 m for geocoding. At the end, the calibrated and corrected images are converted to decibel values.

Sentinel-2 images consist of 13 spectral bands. Each band has a different spatial resolution of 10, 20 or 60 m and represents a different spectral range. Sentinel-2 Level-2A is produced by various processing facilities across Europe using the Sen2Cor and then provided to users through different online platforms (e.g., Copernicus Data Space Ecosystem and Sentinel Hub). Pre-processing procedure is separated in two steps in order to generate ARD. At first, an image classification takes place for the production of the scene classification image, which is useful for masking clouds, thin cirrus and cloud

shadows before processing and analysis or focus analysis on other land covers like water, vegetation and snow. Moreover, the scene classification classifies unsaturated pixels, missing data, dark features and non-vegetated pixels. An amount of pixels remains unclassified for some acquisitions. The second step, is responsible for applying atmospheric corrections to the spectral bands of Sentinel-2 derived images in order to convert top of atmosphere values into surface reflectance. Moreover, final dataset of each image includes supplementary outputs such as aerosol optical thickness, true color image and water vapour. Sentinel-2 Level-2A which is downloaded for CEODC doesn't contain band 10 which is used for atmospheric corrections by Sen2Cor. Bands 2, 3, 4, 8, aerosol optical thickness, true color image and water vapor³ are exported with a spatial resolution of 10 m directly from Sen2Cor. The rest of the bands (i.e., 1, 5, 6, 7, 8A, 9, 11, 12 and scene classification) are resampled to 10 meters spatial resolution using bi-linear interpolation method from Geospatial Data Abstraction Library (GDAL/OGR contributors, 2024).

All of the satellite acquisitions are reprojected to EPSG:32636 before being considered as ARD for spatial alignment and achieve consistency during analysis or fusion of data from different satellite sources. The selected coordinate reference system is considered as the native for Cyprus by EPSG Geodetic Parameter Dataset. Moreover, all the images are converted to COGs which is critical for faster loading and analysis in a cloud-based analytical framework from users. Finally, each data file is converted to a Cloud Optimized GeoTIFF.

For the development of CEODC, the Open Data Cube framework was utilized. Each satellite product according to the ODC framework is considered as product and each acquisition (also what is referred as image) is considered as a dataset. Datasets are organized hierarchically under the corresponding product. CEODC is managing and handling 16 different products. Individual products' definitions according to the eo3 metadata type are generated in YAML format and are used in the ODC to indicate the structure and the format of each dataset. Each product definition contains information like the name of the product, product description, metadata type and the layers of each dataset with relevant information as the name, the data type, the no data value and the measurement units. The next step is the preparation of each dataset's definition. Datasets' definitions are generated according to eo3 metadata type, too. Those contain information like a unique id for each dataset, the eo3 schema, the corresponding product name, the coordinate reference system, the shape and transformation of the grid, the path for each layer's COG in the storage directory, and EO properties like the platform, instrument, date and time of acquisition, file format, a region code, dataset's maturity indication and the product's family. Additionally for Sentinel-2, dataset definition contains cloud coverage on each image to enhance images filtering from users' side. At first, each product is indexed in the data cube using ODC's command line interface and then each dataset is indexed in a similar way. The whole process of preparing definitions and adding products and datasets is an automated pipeline. In addition, the CEODC is built with a microservices architecture on an in-house cloud infrastructure as illustrated in Figure 3. The data storage layer is responsible for keeping ARD rasters and the database where metadata of each geospatial dataset are stored. The microservices layer is responsible for orchestrating the services of data acquisition, raster pre-processing, product and dataset definitions preparation, and finally the indexing to

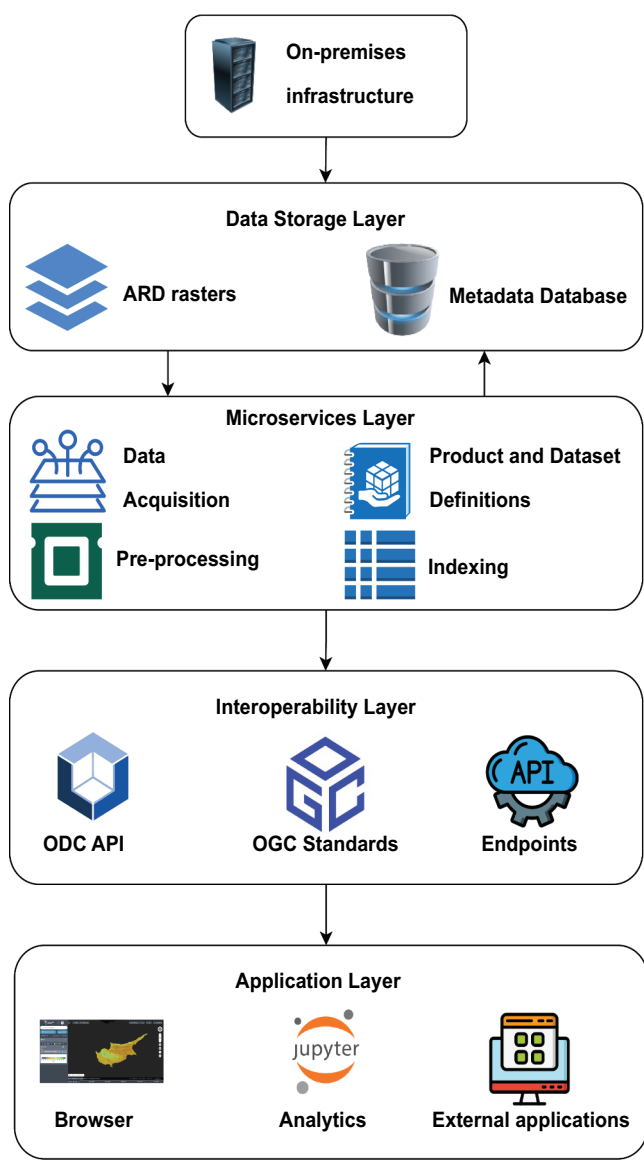


Figure 3. Architecture of CEODC.

the database. Moreover, the interoperability layer is managing all the APIs ODC, and Open Geospatial Consortium (OGC) Standards as also any other endpoints needed to serve data to users, front-end applications, or support downstream applications. Finally, the application layer is managing any front-end application connected to CEODC.

CEODC’s archive is updated daily with the newest available acquisitions by searching the repositories of each aforementioned provider. At the time of writing, it contained 3539, 4845, and 24,578 scenes of Sentinel-1, Sentinel-2, and MODIS, respectively, corresponding to a total volume of 16 TB.

4. Semantic-based water resources monitoring

4.1. Study area

Cyprus is an island situated in the eastern Mediterranean Sea (35°N, 33°E) and is part of the wider area of the Eastern Mediterranean, Middle East, and North Africa region (Eliades et al., 2023). The island covers a total area of 9,254 km² and the climate is characterized as semi-arid with frequently occurring droughts. In general, Cyprus experiences hot dry summers and mild rainy winters which are typical characteristics of the Mediterranean region (Hadjinicolaou et al., 2011). Between years 1991 to 2020, the average mean surface temperature in Cyprus was 11.2 degrees Celsius for the winter season, 16.82 degrees Celsius for the spring season, 26.69 degrees Celsius for the summer season, and 21.03 degrees Celsius for the fall season, according to the reports of the climate crisis in Knowledge Portal of World Bank for 2021⁴. Regarding precipitation amounts, the winter season had 281.06 mm, spring had 90.9 mm, summer had 10.3 mm, and autumn had 87.81 mm.

In this proof-of-concept we focus on the area covered by the structure of Kouris water reservoir as shown in Figure 4 which is located approximately 15 kilometers northwest of Limassol at an altitude of around 250 meters above sea level. Kouris is an artificial structure, constructed in 1988 and is the largest dam among Cyprus' network of 107 dams. The reservoir is fed by Kouris, Limnatis and Kryos rivers. In addition, the dam is connected to Diarizos river via a 14.5 km underground tunnel. The total catchment area covers 300 km² with a water storage capacity of 115 million cubic meters. Finally, the Kouris water reservoir is considered as the most critical water supply infrastructure of the island and thus systematic monitoring is essential to sustain water security of Cyprus.

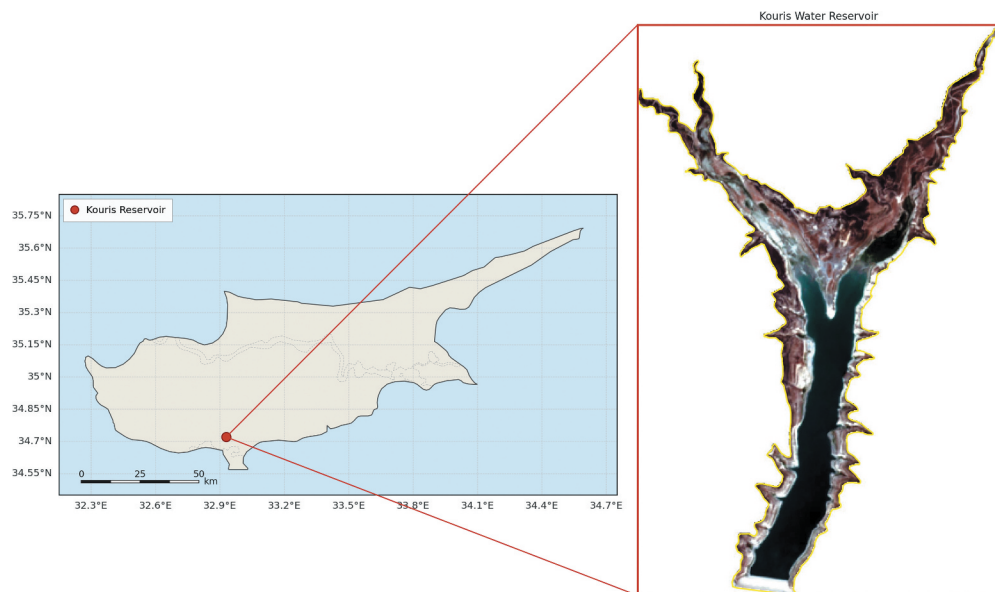


Figure 4. Location of Kouris water reservoir in Cyprus. The red circle indicates the reservoir location on the main map. The inset shows a cloud-free RGB composite from ESA's Sentinel-2 during summer 2025.

4.2. *In-situ data*

Three complementary in-situ datasets were collected to validate satellite-derived water area measurements and enhance the hydrological analysis for Kouris water reservoir. Water storage and inflow data were obtained from the Cyprus Water Development Department at daily and monthly temporal resolution, respectively. Both of the variables are measured in million cubic meters. In-situ dataset of precipitation and water storage temporally covers from 2015 to 2024 while dataset of water inflow is missing data from 2015–2017. In addition, daily precipitation data from the nearest automated weather station of Department of Meteorology of Cyprus were acquired, measured in millimeters. To ensure temporal harmonization with the other two in-situ variables and enable proper correlation analysis, daily precipitation values were aggregated to monthly totals using summation while mean monthly values from water storage data were extracted. The fusion of in-situ data with EO data is taking place within the EODC environment but they are not stored as geospatial products in the EODC. In-situ data can be uploaded in the analytics environment by users according to their needs.

4.3. *Semantic classification framework*

The proposed semantic classification framework translates Sentinel-2 SCL pixel values into six meaningful land cover concepts (i.e., water, vegetation, non-vegetation, clouds, snow, other). We refer to this as a spatially generalized framework because it relies exclusively on the globally consistent SCL, without any site-specific training data or local parameter tuning. As a result, the same semantic rules can be directly applied to other reservoirs and regions where Sentinel-2 L2A data are available. This approach was chosen for several critical reasons.

First, the SCL offers a standardized, analysis-ready classification layer generated by the official ESA's Sen2Cor processor, ensuring both reproducibility and consistency across scenes and time. In addition, the proposed method eliminates subjectivity to site-specific spectral thresholding or manual rule sets that require regional tuning. Thus, SCL with its immediate availability and validation by the Sentinel mission supports automatic large-scale applications and reproducibility in diverse environments and seasons.

Second, this semantic-based framework provides direct interpretability and operation relevance for environmental monitoring with six designated classes. These specific classes are selected by the practical needs of water resource management. "Water", "vegetation", and "non-vegetation" are essential for understanding the water reservoir's surface dynamics, exposed margins and ecosystem response. "Clouds" and "snow" are included because their presence alters the captured scene, while "other" class captures missing data.

Third, compared to modern artificial intelligence (AI) methodologies this approach does not require site-specific labels or manual acquisition of training data. AI-based methods require large high-quality, annotated, and regional or site-specific datasets, and typically result in complex black-box decision processes. In contrast, the proposed framework is fully automated, requires no site-specific training or manual labeling, and provides transparent decision logic and consistency in the outputs.

Table 2 shows semantic land cover concepts, corresponding SCL codes and descriptions. This design enables rapid production of reliable semantic masks, quantitative analysis, time-series trend extraction, and actionable water management indicators.

For each Sentinel-2 image at time t , a binary mask mask_c^t is generated for every semantic concept c with the following equation:

$$\text{mask}_c^t(i, j) = \begin{cases} 1, & \text{if } \text{SCL}^t(i, j) \in \text{codes}(c) \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where (i, j) are the pixel coordinates and $\text{codes}(c)$ is the set of SCL codes associated with concept c as described in **Table 2**.

The total number of pixels N_c^t for semantic concept c at time t is defined as follows:

$$N_c^t = \sum_{ij} \text{mask}_c^t(i, j) \quad (2)$$

Those boolean masks enable automated calculation of quantitative metrics, essential for water resource monitoring. For each semantic concept c and satellite acquisition at time t , spatial coverage A_c^t is computed as:

$$A_c^t = N_c^t \times \alpha, \quad (3)$$

where N_c^t is defined in Eq. 2 and α is a constant equal to 0.01 hectares denoting the unit pixel area of 10-meter resolution. The percentage composition P_c^t is derived as:

$$P_c^t = \frac{N_c^t}{N_{\text{valid}}^t} \times 100, \quad (4)$$

where N_{valid}^t is the number of valid pixels in the acquisition at time t .

4.4. Trend and correlation analysis

In this demonstration study for water resource monitoring through CEODC, a combination of statistical tools is used to identify trends in water availability and quantify correlations between satellite-derived water area and in-situ variables (i.e., water storage, inflow, and precipitation). First, for the correlation analysis, the Pearson correlation coefficient is calculated with the following equation:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (5)$$

Table 2. Semantic land cover concepts and their corresponding Sentinel-2 SCL codes.

Semantic Concept	SCL Code(s)	Description
Water	6	Water bodies
Vegetation	4	Green vegetation
Non-vegetation	5	Bare soil, rocks, and concrete structures
Clouds	3, 8, 9, 10	Cloud shadow, medium and high probability clouds, and cirrus
Snow	11	Snow or ice
Other	0, 1, 2, 7	No data, saturated or defective pixels, shadows, and unclassified

where r is the Pearson correlation coefficient, x_i and y_i represents each paired observation between water area and each given in-situ variable, $\bar{x}$ and $\bar{y}$ denotes the mean values of samples and n is the number of paired observations.

In addition, Spearman rank correlation coefficient was calculated with the following equation:

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}, \quad (6)$$

where ρ is the Spearman rank correlation coefficient, d_i is the difference between the ranks of each pair of observations (x_i, y_i) , and n is the number of paired observations.

Pearson correlation coefficient tests data for linearity under normality assumptions while Spearman rank correlation coefficient detects monotonic relationships without requiring normality. Moreover, in combination the two methods ensure that correlations are not distorted when the relationship is not linear or influenced by outliers. In addition, this combination acts as a cross validation of results where the agreement between the two computed coefficients increases confidence, while differences show if relation is strictly linear or monotonic.

Long-term trends in water availability were analyzed using ordinary least squares regression over the 9-year study period as follows:

$$y = \beta_0 + \beta_1 t + \varepsilon, \quad (7)$$

where y represents the mean annual water area as the arithmetic mean of monthly mean water area values within each hydrological year, t is the hydrological year (e.g., 2016, 2017, 2018, etc.), β_0 is the intercept, β_1 is the trend slope defined as hectares of water area per hydrological year and ε is the error term.

In addition, to ordinary least squares, the non-parametric Mann-Kendall test is employed to assess monotonic trends in water availability without linearity assumptions, making it particularly suitable for hydrological and environmental assessments. Mann-Kendall statistic is calculated as follows:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sgn}(x_j - x_i), \quad (8)$$

where x_i and x_j is water area at hydrological years i and j respectively, and the sign function is given by:

$$\text{sgn}(x_j - x_i) = \begin{cases} +1, & \text{if } x_j - x_i > 0 \\ 0, & \text{if } x_j - x_i = 0 \\ -1, & \text{if } x_j - x_i < 0 \end{cases}$$

The normalized Mann-Kendall statistic τ is computed as:

$$\tau = \frac{S}{\frac{1}{2}n(n-1)}, \quad (9)$$

where n is the number of observations.

Mann-Kendall can only indicate the presence and direction of a trend. Thus, to quantify the magnitude, Sen's slope is estimated with the following equation:

$$Q = \text{median}\left(\frac{x_j - x_i}{j - i}\right), \quad \forall 1 \leq i < j \leq k, \quad (10)$$

where Q represents the estimate of slope (i.e., hectares of water area per hydrological year) and k is the number of hydrological years.

Combining the three methods (i.e., ordinary least squares, Mann–Kendall, and Sen's slope) strengthens the analysis. Agreement between parametric and non-parametric approaches increases confidence while disagreement highlights sensitivity to data assumptions. Furthermore, the utilization of methodologies like Sen's slope can quantify trends of essential environmental variables like water area and availability.

Finally, to determine the water security of the reservoir we used the coefficient of variation for each hydrological year, which is calculated as follows:

$$CV = \frac{\sigma}{\mu}, \quad (11)$$

where CV is the coefficient of variation, σ is the standard deviation of mean monthly water area values within a hydrological year and μ is the mean annual water area of a - hydrological year. Coefficient of variation values for each hydrological year are classified according to demonstrative thresholds as shown in Table 3. Additionally, water stress conditions were determined using the 25th percentile of all water area observations as the threshold for low water availability.

Finally, we determine seasonal timing patterns by identifying peak and minimum water area months for each hydrological year. Inter-annual variability in seasonal timing was quantified using standard deviation of peak and low water months across the 9 hydrological years.

4.5. Experimental setup

Demonstration study of Cyprus Earth Observation Data Cube comprises two datasets from Sentinel-2 acquisitions over Kouris water reservoir at 10-meter spatial resolution. The first dataset consists of 581 satellite observations, used for semantic land cover analysis, correlation analysis between satellite-derived water area and ground-truth data and temporal comparison analysis. In addition, a 15-observation window is applied for the semantic land cover analysis to smooth temporal variations while preserving seasonal patterns.

The second dataset, used for trend analysis, consists of 353 satellite observations containing all of the images where cloud coverage is less than 15% to avoid cloud-induced bias in long-term trend detections and water security assessment.

Table 3. Water availability variability classification based on the coefficient of variation for each hydrological year.

Class	CV Threshold
Stable	< 0.3
Moderate	$0.3 \leq CV < 0.5$
High	≥ 0.5

Both of the datasets are spanning nine hydrological years from 2016 to 2024, where each hydrological year is defined as the period from 1st of October to 30th of September.

All the statistical analyses implemented in Python programming language using the *scipy* and *pymannkendall* packages.

5. Results

5.1. Water resources monitoring

5.1.1. Semantic land cover analysis

Figure 5 shows the semantic land cover dynamics over Kouris water reservoir from 2016 to 2024. The first visual exploration of the figure can easily provide insights on the trade-offs between the different semantic land classes. To provide robust characterization of the composition in the area covered by the reservoir to reduce influence of outliers and skewness, we report median coverage percentages and interquartile ranges (IQR) between 25th and 75th quantile. This approach better captures typical conditions and seasonal patterns compared to mean and standard deviations. Water is covering 25.1% with an IQR of 34.5%, reflecting the typical water extent and its inter-annual variability within the reservoir. Vegetation cover is consistently low, with a median coverage of 2.7% and an IQR of 3.8%. On the other hand non-vegetated surfaces correspond to exposed reservoir beds with a median coverage of 23.3% and an IQR of 39.4%. In addition, presence of clouds accounts for the 7.1% with a wide IQR of 52.7%. Finally, snow cover does not present at all according to this analysis. In addition, a seasonal variation of mean

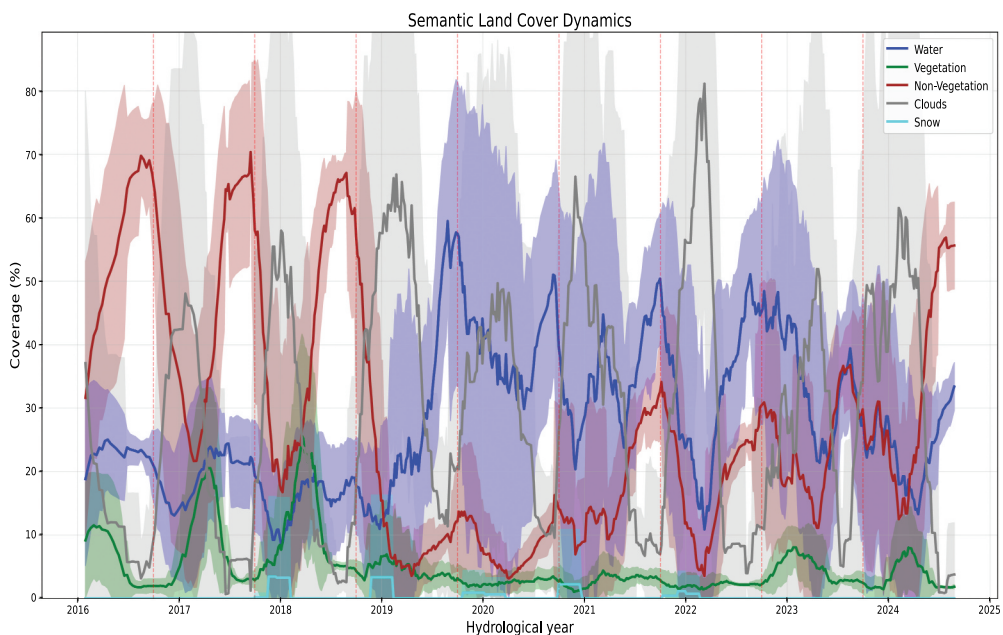


Figure 5. Semantic land cover dynamics. The y-axis illustrates the percentage of area coverage per semantic class and x-axis shows hydrological years. Legend on top right corner presents the corresponding color for each semantic class. Shaded colors denote the standard deviation of area coverage.

water coverage is observed with an average of 32% during summer (i.e., June, July, and August) and autumn months (i.e., September, October, and November), higher than 25.6% and 24.3% observed during winter (i.e., December, January, and February) and spring (i.e., March, April, and May), respectively. Vegetation coverage remained low consistently during all of the seasons, while snow events are minimal averaging 0.3% with high variability likely due to classification uncertainties. Cloud coverage averages a 29.3% with large standard deviation, particularly during winter months reflecting episodic cloud events.

5.1.2. Time-series correlation analysis

Figure 6 and Table 4 show the Pearson and Spearman correlation analysis between satellite-derived water extent and ground-truth variables. In addition, time-series of each ground-truth variable is plotted against the satellite water area to understand temporal dynamics and relationships. Figure 6 highlights the following findings. The strongest correlation is identified between satellite-derived water area with water storage of Kouris reservoir, achieving an r of 0.611 and ρ of 0.626. The nearly equal values indicate a robust linear relationship with minimal influence from outlier or non-normal distributions.

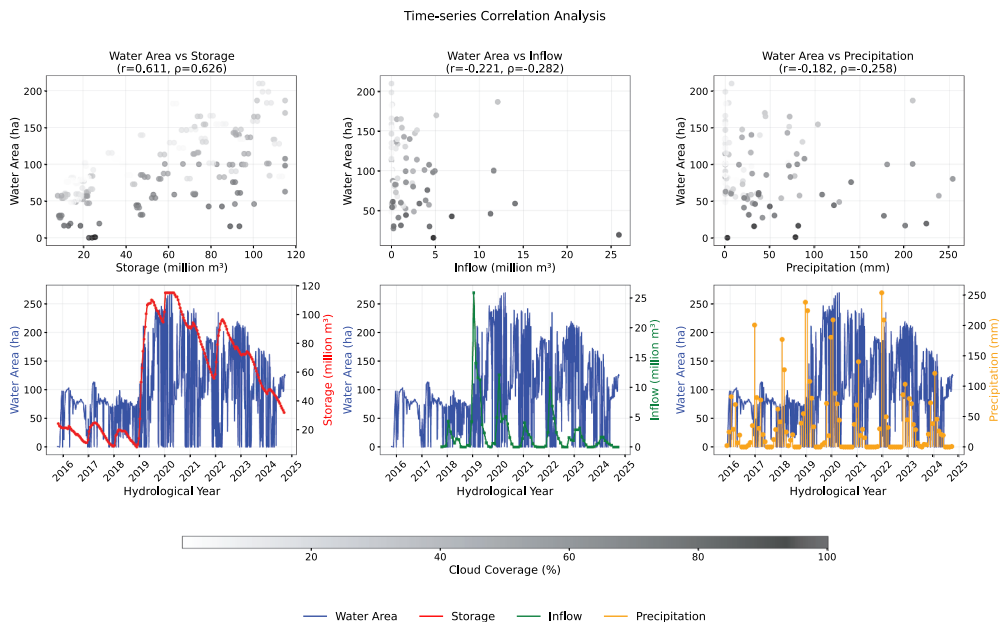


Figure 6. Time-series correlation analysis between satellite-derived water area and in-situ measurements. Top row shows Pearson correlation scatter plots with correlation coefficients (r) for storage, inflow, and precipitation relationships. Middle row presents Spearman rank correlation scatter plots with correlation coefficients (ρ) for the same variables. Bottom row displays temporal comparison time series where the y-axis represents water area (hectares) for satellite data and corresponding units (million cubic meters for storage and inflow, millimeters for precipitation) for ground truth measurements on secondary y-axes. X-axis shows hydrological years (2016–2024). Point shading in scatter plots indicates cloud coverage percentage as shown in the horizontal colorbar. Legend at bottom illustrates colors for satellite-derived water area, storage, inflow, and precipitation.

Table 4. Correlation analysis between water area and in-situ hydrological variables.

Variable	Pearson		Spearman	
	r	p-value	ρ	p-value
Storage	0.611	1×10^{-15}	0.626	1×10^{-15}
Inflow	−0.221	0.004	−0.282	2.14×10^{-4}
Precipitation	−0.182	0.007	−0.258	1.24×10^{-4}

By contrast, inflow and precipitation exhibit weak negative correlation with satellite water extent. Inflow achieves an r of −0.221 and ρ of −0.282 and precipitation an r of −0.182 and ρ of −0.258. The more negative Spearman correlation values suggest monotonic relationships that may not be captured by linear correlation.

The time-series comparison further reveals clear seasonal patterns in satellite-derived water extent that align with expected Mediterranean hydrological cycles. Water storage measurements by Water Development Department of Cyprus are aligned with satellite observations, particularly, during the wet period between 2020–2022, when high inflows sustained large reservoir extents. On the other hand, temporal comparison with inflow and precipitation show weaker alignment, with inflow not directly connected to water extent and precipitation spikes not consistently followed by increases in water extent.

5.1.3. Water security and availability

Figure 7 shows trend analysis of water extent in Kouris reservoir across the 9 hydrological years.

The nine year time-series shows noticeable inter-annual variability in water extent, ranging from a mean of 64.2 hectares during dry year of 2018 to 174.9 hectares in 2020. The monthly mean water extent ranged from an extreme low of 29.9 hectares to an extreme high of 238.4 hectares. From the figure, one can understand that there is a positive trend in mean annual water area. This argument is supported by OLS regression, which estimates an increase of 7.36 hectares per year with p -value lower than 0.001. However, the Mann–Kendall test combined with Sen's slope, which is more robust to linear assumptions and outliers influence, found an increase of 8.86 hectares per year with results found as statistically significant. The substantial difference between minimum and maximum water levels with 29.9 hectares in 2018 and 238.4 hectares in 2020 indicates reservoir's sensitivity to hydrological variability, typical of Mediterranean systems.

Typical peak water month in Mediterranean countries is March where intensive precipitation events come to an end, while typical low water month is September, where most of the irrigation practices in agriculture during summer period have ended. The analysis shows that peak water months are characterized by an inter-annual variation, occurring predominantly in late winter to early spring (i.e., February, March, April). Low water months displayed even greater variability, ranging from February to October across different hydrological years.

The CV analysis revealed that six of the nine hydrological years are characterized by moderate variability, with only three years (i.e., 2016, 2017, and 2020) classified as stable. None of the nine hydrological years exhibited high variability. Both trend analysis methods detected slight increases in variability over time with

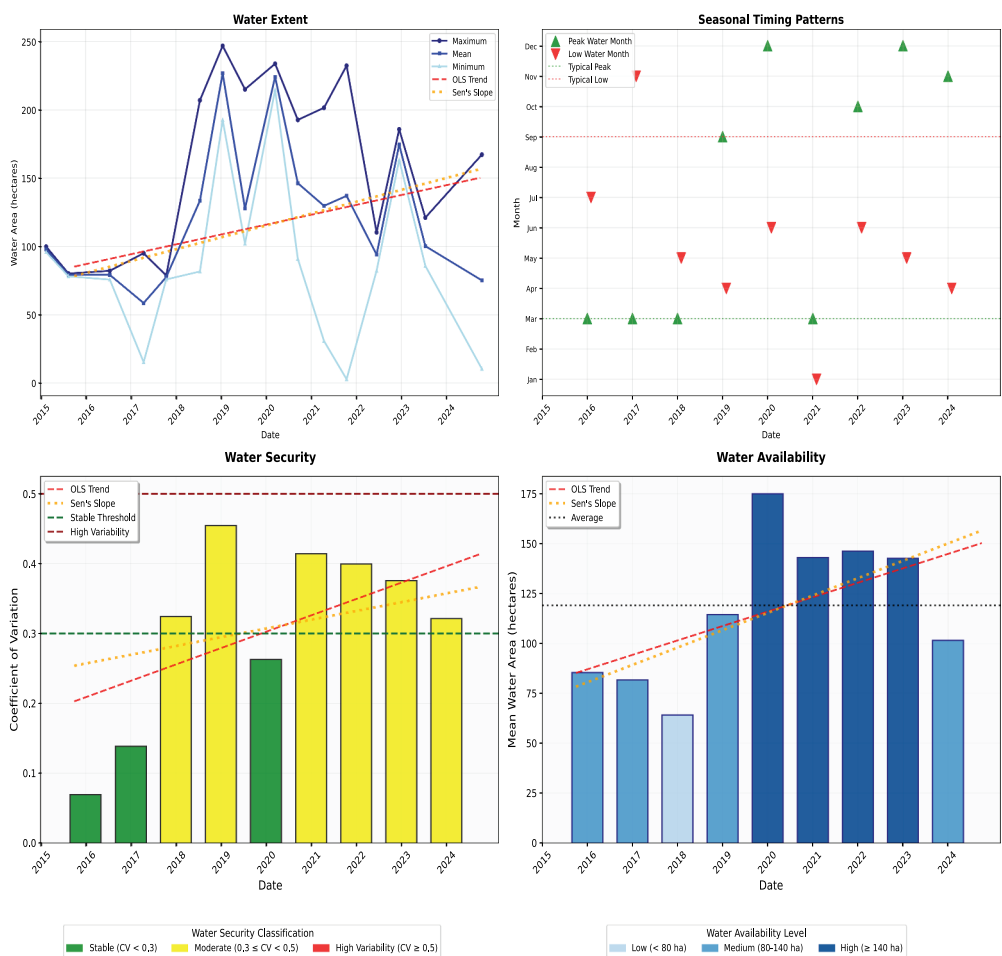


Figure 7. Nine-year hydrological trend analysis. Top row shows annual water extent variability (left) with maximum, mean, and minimum water area trends, and seasonal timing patterns (right) displaying peak and low water months across hydrological years. Bottom row presents water security assessment (left) using coefficient of variation with stability thresholds and trend lines, and long-term water availability (right) showing mean annual water area with trend analysis. Red dashed lines represent OLS regression trends with p -values, dotted orange lines show Sen's slope trends with Mann-Kendall p -values. Bar colors in water security plot indicate stability classification. Bar colors in water availability plot represent water level categories based on absolute thresholds. Legend identifies trend methods and reference lines for each subplot.

0.0239 and 0.0128 hectares per year for OLS and Sen's slope, respectively. In addition, both exhibit p -values of 0.059 and 0.199, respectively, showing no statistical significance in variability changes over the years.

The analysis of water extent dynamics from 2016 to 2024 revealed complex inter-annual patterns with statistically significant long-term trends in water abundance, while there is not a statistical significance in variability.

5.2. Cyprus Earth Observation Data Cube interfaces

To enhance user access and interoperability, we developed two complementary interfaces for the CEODC. Together, these interfaces address both the needs of non-expert users seeking intuitive visualizations and advanced users looking for analytical flexibility.

The first interface is the CEODC Browser, a visualization platform as shown in Figure 8, implemented based on open-access TerriaJS framework, developed by Geoscience Australia. The visualization platform is fully connected to CEODC and is continuously updated with images newly added to the EODC. Images are served from CEODC to Browser using Web Map Services of Open Geospatial Consortium.

Through this platform, users can easily visualize Sentinel-1, Sentinel-2 and MODIS acquisitions, compare images, run time-series, get pixel values, and take high-quality snapshots. In addition, users can upload their own georeferenced data for enhanced analysis. Moreover, we developed a suite of spectral indices related to vegetation, water, land, urban, and natural disasters from Sentinel-2 to enhance visual interpretation of environment by users. The platform is available through <https://browser.cyprusdatacube.com/>.

The second interface is CEODC Analytics, a platform based on JupyterHub as shown in Figure 9 where users can create their own account and apply their own processing and analysis using Python programming language on satellite imagery indexed in Cyprus



Figure 8. Cyprus Earth Observation Data Cube browser. Example screenshot from the visualization platform illustrating a Normalized Difference Water Index image over Kouris water reservoir.

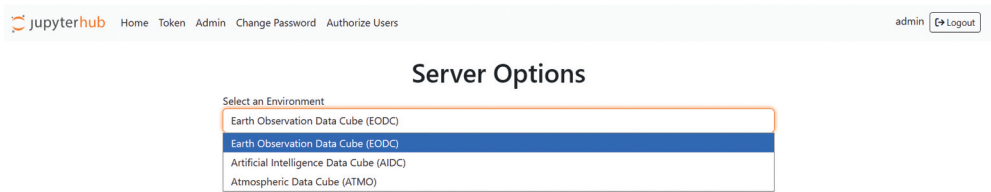


Figure 9. Cyprus Earth Observation Data Cube analytics. Example screenshot from the working environment selection.

Earth Observation Data Cube. Processing resources are allocated to each user on the hosting infrastructure of CEODC. The analytics platform offers the capabilities of three working pre-configured environments with Python packages further than datacube, auxiliary packages (e.g., odc-stac, odc-geo, etc.) and its prerequisites. The first environment, entitled Earth Observation Data Cube, is loaded with well-known scientific packages like rasterio, scipy, and folium. The second, as an extension of the first one, is entitled Artificial Intelligence Data Cube and offers capabilities for building machine learning models like pytorch, tensorflow, and scikit-learn. The third environment, entitled Atmospheric Data Cube, is connected with pylarda, an API for connecting to atmospheric in-situ datasets. Through this, users can access data collected by Cyprus Atmospheric Research Observatory. The platform is available through <https://analytics.cyprusdatacube.com/>. In-situ data can be uploaded in the analytics platform by users to enhance analysis according to their needs.

In addition CEODC ecosystem is accessible through its main website <https://cyprusdatacube.com/>, while all tools and services are integrated also within ERATOSTHENES Regional Exploitation Platform <https://rep.eratosthenes.org.cy/>.

6. Discussions

6.1. *Water resources monitoring*

The results of this work demonstrate the applicability of EODCs and semantic concepts with big data analytics to understand water resources in the context of climate crisis in Mediterranean region. This study utilizes EODC methodological frameworks and demonstrates the role of semantic land cover classification in water reservoir monitoring, where traditional frameworks face unique challenges including high seasonal variability, frequent cloud coverage, and complex management policies.

The semantic land cover approach developed in this study contributes to future EODCs application like the discrimination between water and non-water (e.g., bare soil, concrete structures, vegetation, etc.) reservoir bed exposure during drought conditions, identifying trade-offs between semantic classes, effective handling of mixed pixels in reservoir environment, and robust classification despite high temporal variability in water levels. According to analysis 25.1% of reservoir's structure is covered by water with an IQR of 34.5% demonstrating seasonal sensitivity of the proposed approach and inter-annual fluctuations in water availability. Moreover, non-vegetated areas cover 23.2% of the reservoir's area with an IQR of 39.4% indicating temporal changes corresponding to water volume dynamics. While the median coverage of the area with clouds is 7.1%, the high IQR of 52.7% indicates episodic cloud events. Finally, despite some snow events observed in the area, the 0% for median and IQR shows that such events are rare or such observations reflect to false classification.

Correlation analysis using Pearson and Spearman methods showed a strong correlation ($r = 0.611$, $\rho = 0.626$) between satellite-derived water area coverage and in-situ storage, while inflow and precipitation characterized by weak negative relationships (i.e., inflow : $r = -0.221$, $\rho = -0.282$ and precipitation : $r = -0.182$ and $\rho = -0.258$). This validation exhibits the important role the EODCs can play in water extent monitoring. Moreover, the agreement between the two correlation methods validates our semantic-based approach. However, the

weak correlations with inflow and precipitation reveal important methodological insights about the complexity of satellite-based hydrological monitoring. The weak relationships exhibit the complexity of hydrological processes in water reservoirs, where precipitation and inflow do not directly translate to water extent due to different reasons such as interactions with groundwater, infiltration and dam regulations and management practices.

Both statistical methods detected statistically significant increasing trends of 7.36 hectares per year by OLS and 8.86 hectares per year by Mann–Kendall. On the other hand, visual inspection of water extent shows that the trend reflects recovery from low water volumes between 2015 and 2018, followed by high water availability in 2020, and the observed water decrease in recent years. This observation exhibits a capture of drought recovery in the trend rather than improvement in water security, demonstrating how such trend analysis can be misleading when applied to hydrological processes, main characteristic of Mediterranean systems (Zittis et al., 2019). Moreover, the application of both trend detection methods (i.e., OLS and Mann–Kendall) represents a good practice for EODC time-series analysis, providing robust statistical validation that accounts for different data distributions. In addition, the coefficient of variation analysis demonstrated no significant temporal trend in water variability, indicating that while absolute water levels changed, the relative stability of the system remained consistent throughout the study period. This exhibits that extreme inter-annual variability can hide potential underlying trends and how water reservoirs' management practices by local authorities can influence the extraction of hydrological patterns from water extent dynamics. Moreover, this highlights the importance of incorporating domain expertise into EODC-based water assessments (Bolognesi et al., 2018).

Beyond the scientific analysis, employing the CEODC provides several operational benefits for water resources management. The system allows automatic retrieval and fast time-series analysis of the complete Sentinel-2 archive over the reservoir, reducing manual data handling and enabling near-real-time updates as new imagery becomes available. The semantic framework can be executed reproducibly within the CEODC environment, facilitating routine monitoring. In addition, the integration of CEODC with downstream applications supports decision making and the development of early warning systems for water availability also can be applied for drought and flood detections.

The proposed semantic-based approach differs from well-established approaches like WOfS. WOfS is based on Machine Learning and produces multi-decadal water occurrence products primarily derived from Landsat data with extensive validation at continental scale. In contrast, our approach is simpler and built on Sentinel-2 data requiring no training data and offering 10 m spatial resolution. This simplicity allows easy deployment within EODCs environments such as CEODC and straightforward adaptation by water-related local authorities, even when advanced machine learning expertise is not available. While WOfS is better suited for long-term and large-scale historical characterization, our framework is designed for operational water reservoir monitoring and rapid integration with in-situ data and decision support tools.

Finally, this work shows how EODCs and satellite-derived data can enhance the systematic monitoring of United Nations Sustainable Development Goal (SDG) Indicators, particularly indicators for SDGs 6 (clean water and sanitation), 13 (climate action), 14 (life below water), and 15 (life on land) (Andries et al., 2019; Hakimdavar et al., 2020).

6.2. Cyprus Earth Observation Data Cube interoperability

EODCs benefit from enhanced interoperability using standards characterized by six main dimensions (views) as follows: semantic, geometry, encoding, interaction, interconnection, and composition (Giuliani et al., 2019; Nativi et al., 2017). Regarding semantic interoperability, CEODC uses standardized NASA (e.g., MOD16A2, MOD13Q1, etc.) and ESA (e.g., GRD, MSI, etc.) product code names ensuring international recognition. Moreover, consistent data structure definitions are followed based on eo3 metadata type which includes information like CRS, measurement units, and no data values among others. In addition, each dataset is indexed in CEODC with metadata like the percentage of cloud cover. With respect to geometry interoperability, all the indexed datasets are spatially standardized and aligned with the EPSG:32636. Moreover, to enhance encoding interoperability of CEODC, all the images are translated to COGs and the storage location of each is stored in a PostGIS database system as indicated by ODC. Moreover, user interfaces of CEODC are compliant with Open Geospatial Consortium Web Services through *datacube-ows* including Web Map Service, Web Coverage Service, and Web Map Tile Service. In addition, an open-access JupyterHub is built on top of CEODC to enable users with access to images indexed to CEODC. The CEODC as aforementioned in Section 3 is developed using ODC with respect to framework's standards, making it able to interoperate with external services. In addition, the whole system is developed in a microservices architecture and serves as the data provider of an irrigation scheduling and management application (Neophytides et al., 2024).

6.3. Limitations

Despite the efficient management of satellite-based EO data through the CEODC, there are still some limitations. The CEODC's development is based on a microservices architecture using Docker. However, the system does not utilize advanced orchestration technology, such as Kubernetes, that enables dynamic resource allocation and optimized load balancing. As a result, the system's scalability in future high workloads is limited at the time. Second, despite pre-processing and harmonization procedures taking place for consistent spatial alignment and projection for all the datasets, edge effect and misalignment may persist, especially in pixel-level analysis when fusing MODIS with Sentinel datasets. Addressing these technical and infrastructural limitations will be essential for future upscaling.

On the other hand, the proposed semantic-based framework for water resource monitoring also has limitations. First, the framework is constrained to the model design and classification accuracy of Sen2Cor which generates the utilized Sentinel-2's SCL. Moreover, the utilization of this layer limits the possibility of tailoring the classification process to site-specific land cover classes, thereby limiting adaptability, in contrast to AI-based approaches which may offer higher flexibility.

6.4. Perspectives

To increase the scalability of CEODC, plans for the adoption of an advanced orchestration framework are ongoing. With this, opportunities on interoperability enhancement will

take place by connecting CEODC with regional and global data DCs, enabling the involvement in EO initiatives in the broader area of Eastern Mediterranean, Middle East, and North Africa region. Moreover, the users of CEODC, but also connected downstream applications can benefit from extended data availability from other space-based sensors like Sentinel-3 and Landsat.

Regarding the water resources monitoring framework, we explore the integration of current methods with automated AI-based systems to enhance the daily activities of local authorities and support water-related actions in Cyprus. The future combination of such tasks with AI can enhance semantic-based analysis and information extraction of other meaningful insights for the environment. Moreover, such automated frameworks can be designed for other important land features of the island like forests, urban environments, and agricultural areas.

Beyond the utilization of Sentinel-2 in the case study presented here, other datasets indexed in CEODC open additional avenues for environmental monitoring in general. For example MODIS Evapotranspiration can be used for climate-smart agricultural applications or combined with leaf area index and fraction of photosynthetically active radiation products to assess catchment-scale water and vegetation responses. In addition, those satellite-based products can be used for forest growth monitoring. MODIS Burned Area and Thermal Anomalies products support wildfire monitoring and post-fire impact analysis on environment. Sentinel-1 SAR time-series can be used for soil moisture estimations, disaster risk reduction, and cloud-free environmental monitoring in general, while MODIS Land Surface Products enable the study of drought and heatwave impacts. These potential use cases illustrate that CEODC is not limited to a single sensor or application, but forms the basis for a broader environmental monitoring and security framework.

7. Conclusion

This study showed how EODCs can be used to develop spatially generalized semantic-based frameworks for water resources monitoring to enhance policy and decision making in water scarce areas such as those in the Mediterranean context. In addition, to the best of our knowledge in this study the first operational EODC in the Mediterranean region is presented which provides two user-friendly interfaces (i.e., browser and analytics) with enhanced interoperability to further support environmental monitoring. While the concept of EODCs in general is not new, in this work the integration of CEODC in a water monitoring framework is demonstrated to increase awareness of how EODCs can support environmental monitoring in regions threatened by climate change and therefore aid decision-making procedures of local authorities.

The demonstration case study reveals a strong correlation between satellite-derived water area and in-situ water storage was identified with a Pearson correlation value of 0.611 and Spearman correlation of 0.626 which validates the proposed approach. Furthermore, the semantic classification based on Sentinel-2 SCL product effectively distinguished water from non-water classes across extreme conditions. In addition, the method proved robust across 9 years of Mediterranean hydrological variability. The results revealed complex temporal patterns distinguishing drought recovery from sustained trends and demonstrated how statistical significance requires contextual and domain-based interpretation.

The proposed approach depends on the reliability of Sentinel-2 SCL accuracy for distinguishing water from other semantic classes. Our study consists of a 9 hydrological year dataset. Nevertheless, the utilization of global satellite products enhances the reproducibility and transferability of such frameworks, making it applicable to diverse geographical areas with similar needs.

Such satellite-based hydrological analysis can benefit from longer time-series that can be found in other satellite systems like Landsat which systematically covers world from 1985. In our study, we are demonstrating the framework with the main reservoir of Cyprus as a case study. The demonstration of this framework can benefit from application to other reservoirs in Cyprus and the rest of Mediterranean in the future. Furthermore, the incomplete dataset of inflow can be considered as a limitation of our study. Additionally, such studies can benefit from the application of more novel technologies such as artificial intelligence, and this is a future outlook for better adaptation to other sensors like Sentinel-1 or to enable the framework to produce new outputs like direct estimation of water storage from satellite data and alerting systems for floods and droughts. Moreover, we are looking into the integration of MODIS satellite products in water resource monitoring frameworks for enhanced monitoring of inland water bodies in Cyprus.

Finally, such frameworks can enhance water security monitoring in water scarce regions, support evidence-based reservoir management decisions and enable early warning systems for drought and flood conditions.

Notes

1. https://sentinels.copernicus.eu/documents/d/sentinel/cope-serco-rp-23-1493_sentineldataaccessannual_report_2022
2. https://lpdaac.usgs.gov/product_search/, visited on 23/07/2024
3. https://docs.digitalearthafrika.org/en/latest/data_specs/Sentinel-2_Level-2A_specs.html
4. <https://climateknowledgeportal.worldbank.org/>

Acknowledgements

The authors acknowledge the 'EXCELSIOR': ERATOSTHENES: Excellence Research Centre for Earth Surveillance and Space-Based Monitoring of the Environment H2020 Widespread Teaming project (www.excelsior2020.eu). The 'EXCELSIOR' project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510, from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development and the Cyprus University of Technology. In addition, authors acknowledge Cyprus Water Development Department and Department of Meteorology of Cyprus for the in-situ data provision.

Disclosure statement

No potential conflict of interest was reported by the author(s).

AI disclosure statement

The authors report that they did not use any Generative Artificial Intelligence tools.

Funding

This research was funded by 'EXCELSIOR' (European Union's Horizon 2020 research and innovation programme), grant number 857510.

Data availability statement

All the remote sensing data used in this study can be accessed through cyprusdatacube.com or rep.eratosthenes.org.cy. All the ground-truth data have to be requested from the respective local authority of Cyprus.

Notes on contributors

Stelios P. Neophytides is a researcher at ERATOSTHENES CoE in the Department of Big Earth Data Analytics. He obtained his BSc in Computer Engineering and Informatics from the Cyprus University of Technology (CUT) in 2019. He holds an MSc in Geoinformatics and Geospatial Technologies from CUT (2021). He is currently a PhD candidate in Big Earth Data Analytics at CUT. He previously worked as a Research Assistant at the CYENS Centre of Excellence (Cyprus), where he gained experience in 3D modelling, augmented reality technologies, and digital cultural heritage. His PhD research focuses on artificial intelligence and Big Earth Data analytics in agricultural water management. His research interests include Earth Observation, AI, and Big Earth Data for applications such as crop yield prediction, land cover classification and forecasting, crop type mapping, and water resource management.

Gregory Giuliani is the Head of the Digital Earth Unit and Swiss Data Cube Project Leader at GRID-Geneva of the United Nations Environment Programme (UNEP) and a Senior Lecturer at the University of Geneva's Institute for Environmental Sciences. He is a geologist and environmental scientist who specializes in Remote Sensing, Geographical Information Systems (GIS) and Spatial Data Infrastructures (SDI). He also works at GRID-Geneva of the United Nations Environment Programme (UNEP) since 2001, where he was previously the focal point for Spatial Data Infrastructure (SDI) and is currently the Head of the Digital Earth Unit. Dr. Giuliani's research focuses on Land Change Science and how Earth observations can be used to monitor and assess environmental changes and support sustainable development.

Charalampos Kontoes holds the position of Research Director at the IAASARS/NOA and leads the Operational Unit BEYOND Center for Earth Observation Research and Satellite Remote Sensing. He received his PhD in Earth Observation holding a grant from EC (JRC/ISPR). He leads a team of 55 researchers with active participation in space-related projects funded by ESA, EC, and International Institutions (WB, EIB). He acts as National Delegate contributing to Space Policy and Program Committees of ESA, EC, COPERNICUS, SST/SSA. He is a member of the Steering Committee of Geohazard Supersites and Natural Laboratories of Group of Earth Observations and the Committee of International Union of Geodesy and Geophysics and is actively involved in the GEO Disaster Risk Reduction and Capacity Building Working Groups. He is the GRSS GEO principal and leader of the Disasters Resilience and Health Action Group of EuroGEO. He is responsible for Node 3 of Copernicus Data Access Service, the Hellenic Mirror Site, and the NOA's Ground Segment. He is the author of more than 200 publications (h-index 35). He speaks English, French and Italian.

Diofantos G. Hadjimitsis is a Full Professor at the Department of Civil Engineering and Geomatics, Faculty of Engineering and Technology of the Cyprus University of Technology (CUT) and the Managing Director of the 'ERATOSTHENES Centre of Excellence'. He is the former 'Vice-Rector of Academic Affairs' of the Cyprus University of Technology (2016-2019) with main duties of chairing and managing the 'Research', 'Quality Assurance', 'Studies' committees for the whole university. During his term, his management and promotion efforts achieved a high ranking for the University, e.g., Ranked #351-400 in 2018 and 301-350th in 2019-2020 in the Times Higher Education World

University Rankings. He is the Vice President of the 'The Cyprus Agency of Quality Assurance and Accreditation in Higher Education' board and he is a member of the board since 2015. He is included in the world's top 2% of most important scientists in the list issued by Stanford University.

Michalis Mavrovouniotis received the B.Sc. (Hons) degree in Computer Science from the University of Leicester, UK (2008), the M.Sc. degree in Natural Computation from the University of Birmingham, UK (2009), and Ph.D. degree in Computer Science from the University of Leicester, UK (2013). He is the Coordinator of the Big Earth Data Analytics Department of the ERATOSTHENES Centre of Excellence. His research interests include evolutionary computation, neural computation, swarm intelligence, memetic computing, combinatorial optimization problems, heuristic and metaheuristics, artificial intelligence in dynamic and uncertain environments, and relevant real-world applications. He has contributed to several funded projects with partners in both academia and industry. Dr. Mavrovouniotis has more than 80 peer-reviewed publications with h-index of 30 according to Google Scholar. He is currently the Chair of the IEEE Task Force on Evolutionary Computation in Dynamic and Uncertain Environments, a member of the IEEE Computational Intelligence Society and an Associate Fellow of the UK Higher Education Academy.

ORCID

Stelios P. Neophytides  <http://orcid.org/0000-0002-1678-1556>

References

- Ajibola, S., & Cabral, P. (2024). A systematic literature review and bibliometric analysis of semantic segmentation models in land cover mapping. *Remote Sensing*, 16(12), 2222. <https://doi.org/10.3390/rs16122222>
- Allu, A. R., & Mesapam, S. (2025). Impact of remote sensing data fusion on agriculture applications: A review. *The European Journal of Agronomy*, 164, 127478. <https://doi.org/10.1016/j.eja.2024.127478>
- Andries, A., Morse, S., Murphy, R. J., Lynch, J., & Woolliams, E. R. (2019). Seeing sustainability from space: Using Earth observation data to populate the UN sustainable development goal indicators. *Sustainability*, 11(18), 5062. <https://doi.org/10.3390/su11185062>
- Asmeryan, S., Muradyan, V., Tepanosyan, G., Hovsepyan, A., Saghatelian, A., Astsatryan, H., Grigoryan, H., Abrahamyan, R., Guigoz, Y., & Giuliani, G. (2019). Paving the way towards an Armenian data cube. *Data*, 4(3), 117. <https://doi.org/10.3390/data4030117>
- Augustin, H., Sudmanns, M., Tiede, D., Lang, S., & Baraldi, A. (2019). Semantic earth observation data cubes. *Data*, 4(3), 102. <https://doi.org/10.3390/data4030102>
- Augustin, H., Sudmanns, M., Weber, H., Baraldi, A., Wunderle, S., Neuhaus, C., Reichel, S., Van Der Meer, L., Hummer, P., & Tiede, D. (2021, April 19–30). SemantiX: A cross-sensor semantic EO data cube to open and leverage AVHRR time-series and essential climate variables with scientists and the public. *EGU General Assembly 2021*, EGU21–12722. <https://doi.org/10.5194/egusphere-egu21-12722>
- Baumann, P., Dehmel, A., Furtado, P., Ritsch, R., & Widmann, N. (1998). The multidimensional database system Rasdaman. In *Proceedings of the 1998 ACM SIGMOD International Conference on Management of Data, SIGMOD '98*, New York, NY, USA (pp. 575–577). Association for Computing Machinery. <https://doi.org/10.1145/331697.331732>
- Baumann, P., Furtado, P., Ritsch, R., & Widmann, N. (1997). The RasDaMan approach to multidimensional database management. In *Proceedings of the 1997 ACM symposium on Applied computing - SAC '97* (pp. 166–173). ACM Press, San Jose, California, United States.
- Bolognesi, T., Gerlak, A. K., & Giuliani, G. (2018). Explaining and measuring social-ecological pathways: The case of global changes and water security. *Sustainability*, 10(12), 4378. <https://doi.org/10.3390/su10124378>

- Câmara, G., Vinhas, L., Ferreira, K. R., Queiroz, G. R. D., Souza, R. C. M. D., Monteiro, A. M. V., Carvalho, M. T. D., Casanova, M. A., & Freitas, U. M. D. (2008). Terralib: An open source GIS library for large-scale environmental and socio-economic applications. In G. B. Hall & M. G. Leahy (Eds.), *Open source approaches in spatial data handling* (pp. 247–270). Springer.
- Cao, Q., Li, G., Yao, X., & Ma, Y. (2022). China data cube (CDC) for big Earth observation data: Practices and lessons learned. *Information*, 13(9), 407. <https://doi.org/10.3390/info13090407>
- Chatenoux, B., Richard, J.-P., Small, D., Roeoesli, C., Wingate, V., Poussin, C., Rodila, D., Peduzzi, P., Steinmeier, C., Ginzler, C., Psomas, A., Schaepman, M. E., & Giuliani, G. (2021). The Swiss data cube, analysis ready data archive using earth observations of Switzerland. *Scientific Data*, 8(1), 295. <https://doi.org/10.1038/s41597-021-01076-6>
- Chaves, M. E. D., Soares, A. R., Sanches, I. D., & Fronza, J. G. (2021). Cbers data cubes for land use and land cover mapping in the Brazilian Cerrado agricultural belt. *International Journal of Remote Sensing*, 42(21), 8398–8432. <https://doi.org/10.1080/01431161.2021.1978584>
- Claro, A. M., Fonseca, A., Fraga, H., & Santos, J. A. (2024). Future agricultural water availability in Mediterranean countries under climate change: A systematic review. *Water*, 16(17), 2484. <https://doi.org/10.3390/w16172484>
- CNES and contributors, E. (2023). *Eodag: Earth observation data access gateway*. <https://github.com/CS-SI/eodag>
- Dhu, T., Giuliani, G., Juárez, J., Kavvada, A., Killough, B., Merodio, P., Minchin, S., & Ramage, S. (2019). National open data cubes and their contribution to country-level development policies and practices. *Data*, 4(4), 144. <https://doi.org/10.3390/data4040144>
- Drivas, T., Sitokonstantinou, V., Tsardanidis, I., Koukos, A., Kontoes, C., & Karathanassi, V. (2022). A data cube of big satellite image time-series for agriculture monitoring. In *2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)* (pp. 1–5). IEEE. <https://doi.org/10.1109/IVMSP54334.2022.9816291>
- Eliades, M., Michaelides, S., Evagorou, E., Fotiou, K., Fragkos, K., Leventis, G., Theocharidis, C., Panagiotou, C. F., Mavrovouniotis, M., Neophytides, S., Papoutsas, C., Neocleous, K., Themistocleous, K., Anayiotos, A., Komodromos, G., Schreier, G., Kontoes, C., & Hadjimitsis, D. (2023). Earth observation in the EMMENA region: Scoping review of current applications and knowledge gaps. *Remote Sensing*, 15(17), 4202. <https://doi.org/10.3390/rs15174202>
- European Space Agency. (2023). *Snappy: Python API for Snap*. <https://step.esa.int/main/toolboxes/snap/python-snappy/>
- Gans, F., Mahecha, M. D., Reichstein, M., Brandt, G., Fomferra, N., Permana, H., & Brockmann, C. (2018). The earth in a box: A light-weight data cube approach to empower the study of land-surface processes and interactions. In *EGU General Assembly Conference Abstracts. Geophysical Research Abstracts, EGU General Assembly 2018* (p. 9841).
- GDAL/OGR contributors. (2024). *GDAL/OGR geospatial data abstraction software library*. Open Source Geospatial Foundation. <https://gdal.org>
- Giuliani, G., Masó, J., Mazzetti, P., Nativi, S., & Zabala, A. (2019). Paving the way to increased interoperability of Earth observations data cubes. *Data*, 4(3), 113. <https://doi.org/10.3390/data4030113>
- Hadjinicolaou, P., Giannakopoulos, C., Zerefos, C., Lange, M. A., Pashiardis, S., & Lelieveld, J. (2011). Mid-21st century climate and weather extremes in Cyprus as projected by six regional climate models. *Regional Environmental Change*, 11(3), 441–457. <https://doi.org/10.1007/s10113-010-0153-1>
- Hakimdavar, R., Hubbard, A., Policelli, F., Pickens, A., Hansen, M., Fatoyinbo, T., Lagomasino, D., Pahlevan, N., Unninayar, S., Kavvada, A., Carroll, M., Smith, B., Hurwitz, M., Wood, D., & Schollaert Uz, S. (2020). Monitoring water-related ecosystems with Earth observation data in support of Sustainable Development Goal (SDG) 6 reporting. *Remote Sensing*, 12(10), 1634. <https://doi.org/10.3390/rs12101634>
- Halabisky, M., Yuan, F., Adimou, G., Birchall, E., Boamah, E., Burton, C., Chong, E.-F., Hall, L., Jorand, C., Leith, A., Lewis, A., Mamane, B., Mar, F., Moghaddam, N., Ongo, D., & Rebelo, L.-M. (2024). A dynamic surface water extent service for Africa developed through continental-scale

- collaboration. *Frontiers in Environmental Science* 12:1251315 doi:<https://doi.org/10.3389/fenvs.2024.1251315>.
- Hanson, M. (2019). The open-source software ecosystem for leveraging public datasets in spatio-temporal asset catalogs (STAC). In *Volume 2019 of AGU Fall Meeting Abstracts*, pages IN23B-07. AGU.
- Kavouras, I., Rallis, I., Doulamis, N., & Doulamis, A. (2024). Ocean-dc: An analysis ready data cube framework for environmental and climate change monitoring over the port areas. In *Proceedings of the 17th International Conference on Pervasive Technologies Related to Assistive Environments, PETRA '24* (pp. 412–419). Association for Computing Machinery.
- Khan, J., E, S., & Ramesh, R. (2024). Performance evaluation of geospatial images based on Zarr and TIFF. *arXiv preprint arXiv:2411.11291*. <https://doi.org/10.48550/arXiv.2411.11291>
- Killough, B. (2018). Overview of the open data cube initiative. In *IGASS, 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium* (pp. 8629–8632). IEEE. <https://doi.org/10.1109/IGARSS.2018.8517694>
- Killough, B. (2019). The impact of analysis ready data in the Africa regional data cube. In *IGARSS, 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium* (pp. 5646–5649). IEEE, Yokohama, Japan.
- Killough, B., Siqueira, A., & Dyke, G. (2020). Advancements in the open data cube and analysis ready data - past, present and future. In *IGARSS, 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium* (pp. 3373–3375). IEEE. <https://doi.org/10.1109/IGARSS39084.2020.9324712>
- Kreiser, Z., Killough, B., & Rizvi, S. R. (2018). Water across synthetic aperture radar data (WASARD): SAR water body classification for the open data cube. In *IGARSS, 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium* (pp. 437–440). IEEE Xplore. <https://doi.org/10.1109/IGARSS.2018.8517447>
- Kröber, F., Sudmanns, M., Abad, L., & Tiede, D. (2025). On-demand, semantic EO data cubes - knowledge-based, semantic querying of multimodal data for mesoscale analyses anywhere on Earth. *ISPRS Journal of Photogrammetry & Remote Sensing*, 228, 552–565. <https://doi.org/10.1016/j.isprsjprs.2025.07.015>
- Lewis, A., Oliver, S., Lymburner, L., Evans, B., Wyborn, L., Mueller, N., Raevksi, G., Hooke, J., Woodcock, R., Sixsmith, J., Wu, W., Tan, P., Li, F., Killough, B., Minchin, S., Roberts, D., Ayers, D., Bala, B., Dwyer, J. . . Wang, L.-W. (2017). The Australian geoscience data cube -foundations and lessons learned. *Remote Sensing of Environment*, 202, 276–292. <https://doi.org/10.1016/j.rse.2017.03.015>
- López, L., Fisher, M., Friesz, A., Wu, Q., Steiker, A., & Community, E. (2024). *Earthaccess: Accelerating NASA Earthdata access through open, collaborative development*. <https://nasa-openscapes.github.io/news/2024-03-04-earthaccess-tech-spotlight/>
- Malthus, T. J., Lehmann, E., Ho, X., Botha, E., & Anstee, J. (2019). Implementation of a satellite based inland water algal bloom alerting system using analysis ready data. *Remote Sensing*, 11(24), 2954. <https://doi.org/10.3390/rs11242954>
- Mantas, A., Yfantis, F., Bilidas, D., Stamoulis, G., Kondylatos, S., Prapas, I., Papoutsis, I., María Tárraga Habas, J., Sevillano Marco, E., Castel, F., Laine, C., & Koubarakis, M. (2024). Plato: A semantic data cube system using ontology-based data access technologies. *IEEE Access*, 12, 130356–130374. <https://doi.org/10.1109/ACCESS.2024.3453494>
- Montero, D., Kraemer, G., Anghilea, A., Aybar, C., Brandt, G., Camps-Valls, G., Cremer, F., Flik, I., Gans, F., Habershon, S., Ji, C., Kattenborn, T., Martínez-Ferrer, L., Martinuzzi, F., Reinhardt, M., Söchting, M., Teber, K., & Mahecha, M. D. (2024). Earth system data cubes: Avenues for advancing Earth system research. *Environmental Data Science*, 3, e27. <https://doi.org/10.1017/eds.2024.22>
- Mubea, K., Boamah, E., Rebelo, L., Yuan, F., & Lewis, A. (2024). Digital Earth Africa-operational EO-based insights for Africa's development challenges. In J. Potel, K. Labbassi, S. Tesfamichael, H. Annegarn, J. Kufoniya, & S. Wade (Eds.), *Space and geospatial technologies for the Africa we want* (pp. 487–495). Springer Nature.
- Mueller, N., Lewis, A., Roberts, D., Ring, S., Melrose, R., Sixsmith, J., Lymburner, L., McIntyre, A., Tan, P., Curnow, S., & Ip, A. (2016). Water observations from space: Mapping surface water from 25 years

- of Landsat imagery across Australia. *Remote Sensing of Environment*, 174, 341–352. <https://doi.org/10.1016/j.rse.2015.11.003>
- Nativi, S., Mazzetti, P., & Craglia, M. (2017). A view-based model of data-cube to support big earth data systems interoperability. *Big Earth Data*, 1(1–2), 75–99. <https://doi.org/10.1080/20964471.2017.1404232>
- Neophytides, S. P., Eliades, M., Mavrovouniotis, M., Papoutsas, C., Papadavid, G., & Hadjimitsis, D. G. (2024). Improved water resources management for smart farming: A case study for Cyprus. *Scientific Reports*, 14(1), 31751. <https://doi.org/10.1038/s41598-024-82797-7>
- Peppas, M., Solano-Correa, Y., Mills, J., & Haile, A. (2023). Rapid mapping of waterbody variations in the central Rift Valley, Ethiopia, using the Digital Earth Africa open data cube. In *IGARSS, 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium* (pp. 5359–5362). IEEE. <https://doi.org/10.1109/IGARSS52108.2023.10282717>
- Peters, J., Neumann, A., Jaeger, M., Gienapp, L., & Umlauf, J. (2025). Machine learning toolkits for Earth system data cubes. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(27), 28302–28311. [ml4xcube](https://doi.org/10.1109/ML4XCUBE)
- Picoli, M. C. A., Simoes, R., Chaves, M., Santos, L. A., Sanchez, A., Soares, A., Sanches, I. D., Ferreira, K. R., & Queiroz, G. R. (2020). Cbers data cube: A powerful technology for mapping and monitoring Brazilian biomes. *ISPRS Annals of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, V-3-2020, 533–539. <https://doi.org/10.5194/isprs-annals-V-3-2020-533-2020>
- Poussin, C., Peduzzi, P., & Giuliani, G. (2025). Snow observation from space: An approach to improving snow cover detection using four decades of Landsat and Sentinel-2 imageries across Switzerland. *Science of Remote Sensing*, 11, 100182. <https://doi.org/10.1016/j.srs.2024.100182>
- Söchting, M., Mahecha, M. D., Montero, D., & Scheuermann, G. (2024). Lexcube: Interactive visualization of large earth system data cubes. *IEEE Computer Graphics and Applications*, 44(1), 25–37. <https://doi.org/10.1109/MCG.2023.3321989>
- Sun, Y., Wang, D., Li, L., Ning, R., Yu, S., & Gao, N. (2024). Application of remote sensing technology in water quality monitoring: From traditional approaches to artificial intelligence. *Water Research*, 267, 122546. <https://doi.org/10.1016/j.watres.2024.122546>
- Wang, Z., & Gao, Z. (2022). Dynamic monitoring of flood disaster based on remote sensing data cube. *Natural Hazards*, 114(3), 3123–3138. <https://doi.org/10.1007/s11069-022-05508-3>
- Wilkinson, R., Mleczo, M., Brewin, R., Gaston, K., Mueller, M., Shutler, J., Yan, X., & Anderson, K. (2024). Environmental impacts of earth observation data in the constellation and cloud computing era. *Science of the Total Environment*, 909, 168584. <https://doi.org/10.1016/j.scitotenv.2023.168584>
- xcube-dev. (2025). *Xcube: An xarray-based EO data cube toolkit*. <https://github.com/xcube-dev/xcube>
- Zittis, G., Hadjinicolaou, P., Klangidou, M., Proestos, Y., & Lelieveld, J. (2019). A multi-model, multi-scenario, and multi-domain analysis of regional climate projections for the Mediterranean. *Regional Environmental Change*, 19(8), 2621–2635. <https://doi.org/10.1007/s10113-019-01565-w>