Ocean-DC: An analysis ready data cube framework for environmental and climate change monitoring over the port areas

Ioannis Kavouras*
ikavouras@mail.ntua.gr
National Technical University of Athens
9, Iroon Polytechniou st, 15772, Zografou,, Greece

Ioannis Rallis*
irallis@central.ntua.gr
National Technical University of Athens
9, Iroon Polytechniou st, 15772, Zografou,, Greece

Nikolaos Doulamis*
ndoulam@cs.ntua.gr
National Technical University of Athens
9, Iroon Polytechniou st, 15772, Zografou,, Greece

Anastasios Doulamis*
adoulam@cs.ntua.gr
National Technical University of Athens
9, Iroon Polytechniou st, 15772, Zografou,, Greece

ABSTRACT

The environmental hazards and climate change effects causes serious problems in land and coastal areas. A solution to this problem can be the periodic monitoring over critical areas, like coastal region with heavy industrial activity (i.e., ship-buildings) or areas where a disaster (i.e., oil-spill) has occurred. Today there are several Earth and non-Earth Observation data available from several data providers. These data are huge in size and usually it is needed to combine several data from multiple sources (i.e., data with format differences) for a more effective evaluation. For addressing these issues, this work proposes the Ocean-DC framework as a solution in data harmonization and homogenization. A strong advantage of this Data Cube implementation is the generation of a single NetCDF product that contains Earth Observation data of several data types (i.e., Landsat-8 and Sentinel-2). To evaluate the effectiveness and efficiency of the Ocean-DC implementation, it is examined a case study of an oil-spill in Saronic gulf in September of 2017. The generated 4D Data Cube considers both Landsat-8,9 and Sentinel-2 products for a time-series analysis, before, during, and after the oil-spill event. The Ocean-DC framework successfully generated a NetCDF product, containing all the necessary remote sensing products for monitoring the oil-spill disaster in the Saronic gulf.
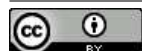
CCS CONCEPTS

• **Information systems** → **Data structures**; *Expert systems*; **Data analytics**; **Geographic information systems**.

KEYWORDS

Data Cubes, Climate Change, Coastal Areas

---

*All authors contributed equally to this research.

1 INTRODUCTION

The modern urban environments have been developed with the purpose of interconnecting people, infrastructures, activities and several other resources, making them susceptible to environmental disasters and climate change effects [18, 22, 26]. Daily industrial activities, especially near the port areas, have a negative impact on the local environment and climate endangering both the human health and well-being, as well as the local nature species.

All the human activities result in the following issues: (a) lack of urban greenness [36]; (b) flooding [28]; (c) non-healthy atmospheric quality [30]; (d) increment of greenhouse gases emissions [32]; and (e) the phenomenon of heat island effect [27]. In most of these cases, the impact can be evaluated by continuously monitoring the activities over the area of interest using Earth Observation (EO) data provided by space-borne technologies (i.e., satellites) and aerial or field climate measurements. In other cases, the impact can be instantaneous, causing a huge disaster over a radius, like in an oil-spill situation.

Today, there is available a large number of EO data and it expands each week [12]. Moreover, the available data are produced from different data providers (i.e., Copernicus, USGS, ESA, NASA, etc.), thus there are differences in distribution format, spatial resolution, data collection, etc. In some cases, however, two missions from different provides (i.e., Landsat-8 and Sentinel-2) have similarities, thus such data can be used in combination for thickening a time-series dataset. For a successful data harmonization and homogenization a Data Cube (DC) implementation can be used [17].

A DC [4, 15] is a multidimensional array containing all the necessary information (i.e., data, metadata, coordinates, etc.) describing a phenomenon. To better understand a basic Earth Observation DC framework, let us denote as 2D array $(B_i(x, y))$ corresponding an spectral band (i.e., image) of a satellite product. Thus, each coordinate set (x, y) corresponds to a radiometric reflectance value for this band. However, each satellite retrieves data on several
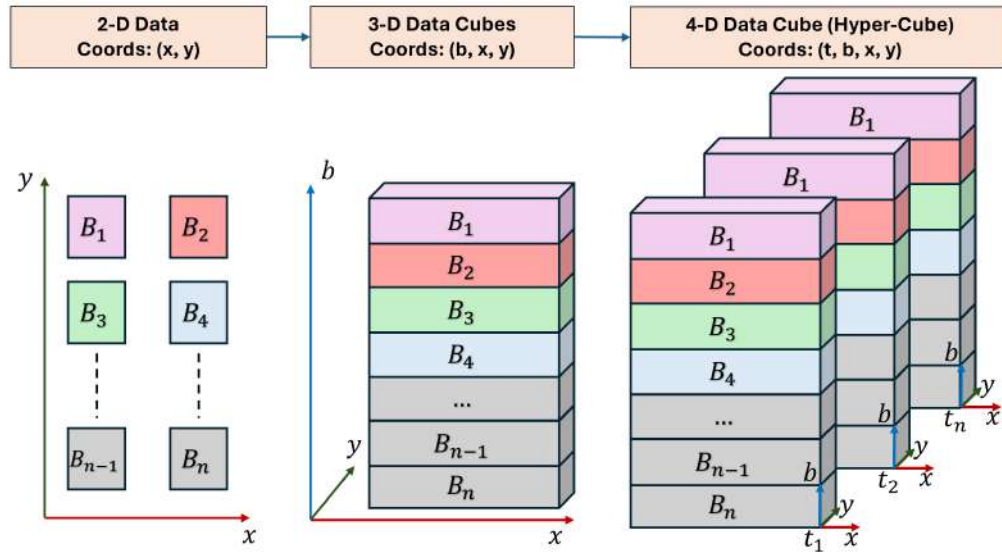
Figure 1: An example of a Data Cube's dimension expanding.

wavelengths, which means that they generate several images for different wavelengths of the spectrum (i.e., Coastal Aerosol, Blue, Green, Red, Near Infrared, etc.). The combination of these images can generate several remote sensing products [38, 41] for indicating a phenomenon. Usually, the raw data are captured in different spatial resolutions (i.e., 10m/px, 15m/px, 30m/px, 60m/px, etc), thus any combinations and/or calculations can not be achieved in the original sate of data. For addressing this issue, let us now denote a 3D array ($DC_i(b, x, y)$) for storing all the spectral bands N ($DC_i(b, x, y) = [B_1(x, y), B_2(x, y), ..., B_N(x, y)]$) in a common spatial resolution format and coordination reference system. Finally, the monitoring of a phenomenon introduces the parameter of time (t). Thus, for effectively monitoring a periodic phenomenon (i.e., climate) over an are and detecting abnormalities it is necessary to generate a DC containing data from various sources. Thus, a final DC implementation can be a 4D-DC (Hypercube) defined as $HDC_i(t, b, x, y) = [DC_1(b, x, y), DC_2(b, x, y), ...DC_N(b, x, y)]$, containing N 3D-DC of different time periods (i.e., before and after an event). Figure 1 illustrates the DC dimension expanding as described in this paragraph.

For the successful implementation of a 4D-DC it is necessary to harmonize and homogenize all of the available 3D-DCs. If the original data are provided from the same data provider, then the implementation can follow the same processing, however additional processes are necessary if the data are provided from more than one sources. For addressing this challenge, in this work, we propose the Oceanids Data Cubes (Ocean-DCs) as a novel approach for continuously and quickly monitoring an area of interest using several kinds of data from the same or different data sources. Moreover, the Ocean-DC provides data harmonization and homogenization by applying the common template for Landsat-8 and Sentinel-2 data products, as well as it offers a complete ready to use analysis products by calculating the most common used remote sensing indices. To be more precise, Ocean-DCs' products can be used as is

in a decision support system platform alongside in synergy with other data of the platform for indicating environmental and climate change hazards in many applications, emphasizing to port, marine and coastal applications (i.e., monitoring, forecasting, etc.).

For demonstrating the Ocean-DC implementation a monitoring case study is provided in this manuscript. For the case study it is selected the oil-spill disaster that took place on the 10th of September 2017 in the Saronikos gulf (Greece) causing huge environmental disaster to the surrounded shores and sea/water area. This case study aims to provide an application example of the proposed Ocean-DC for periodically monitoring a phenomenon based on time-series analysis. However, Ocean-DC products can be used differently depended on the application and the related phenomenon.

The rest of this paper is organized as follow: (a) Section 2 presents the related literature review; (b) Section 3 describes the proposed Ocean-DC framework; (c) Section 4 provides a brief demonstration of environmental monitoring based on an Ocean-DC product; (d) Section 5 discuss the outcomes of this work and (e) Section 6 summarizes this work.

2 RELATED WORK

A Data Cube is a multidimensional array capable or a data structure that allows the efficient analysis of big data and complex data structures along the multiple dimensions. A DC framework is mostly used for applications regarding big data storage and online analytical processing (OLAP) [31] based on complex queries [13] and analysis [9]. Major functionalities of a DC framework are: (a) multidimensional representation [19]; (b) cube structure for big data visualizations [33]; (c) data aggregation [10]; (d) fast query performance [23]; (e) OLAP operations [34]; and (d) smart big data applications [21].

The current DC's implementations are emphasizing on organizing the data on multiple dimensions [14] like time, geography,

product, or customer. Each dimension represents a different aspect of the data. Thus, DCs structure is often visualized as cube with multiple dimensions. Each cell in the cube's representation is a unique combination of dimensional values. For example, in a remote sensing DC an indicative combination can be the (time x spectral bands x geographical coordinates). Moreover, in other DC examples, like climatic and geographical DC, the data can be aggregated using traditional aggregation functions like mean, count, average, minimum, and maximum.

However, it is important to note that the numerical values alone are not enough to describe a phenomenon. Another strong characteristic of the DC's architecture is the inclusion of metadata information that describes the numerical values of the data [5]. Such information includes the metric system of the data, the origin of the data, geo-spatial information, etc. This information can be used for achieving fast query performance, which is one of the main advantages of the utilization of DCs, especially when dealing with large volumes of data (i.e., addressing big data storage performance issues).

Several Data Cube implementations have been suggested the recent years. An indicative example can be the work of Ferreira et al. [8], where they proposed an effectual way for generating Earth Observation Data Cubes (EODCs) over the Brazil region. They additionally are exploiting the benefits of the Open Data Cubes (ODC) platform, contributing by conducting a analysis-ready data and multidimensional data cubes from remote sensing images to effectively mapping the land use and land cover based on employing remote sensing imagery time series analysis and machine learning techniques.

The Australian Geoscience Data Cubes (ADGCs) is one of the first data cubes implementation, proposed by Lewis et al. [20] and manages the challenges of the Big Data, such us the storage volume, velocity, and polymorphism or variety, limiting the earth observation data usefulness. Some of the application, where the ADGCs have been applied are: (a) water observations; (b) coastal erosion; (c) agriculture monitoring; (d) forest cover changes; and (e) biodiversity.

In the work of Poussin et al. [29] the advantages and limitations of the Earth Observation Swiss Data Cubes (EOSDCs) are exploited by monitoring the snow cover change in Gran Paradiso national Park, Alps. EOSDCs architecture is based on the Open Data Cubes software [6], which is a geospatial data management and analysis software based on open-source technologies for processing remote sensing data by providing a framework for accessing, storing, managing, and analyzing grid-aligned satellite Earth Observation data collections in huge quantities.

Other works utilizing the EOSDCs implementations are proposed by Giuliani et al. [10, 11]. In the first work [10] they propose a methodology for generating accurate and consistent land degradation products based on remote sensing imagery datasets. They uses a data cube and several sun-indicators for assessing land productivity and cover, as well as soil organic carbon changes for evaluating land degradation. The second work [11] suggests the usage of a data cube on demand (DCoD), which is a script-based tool, for automating the generation of data cubes under predetermined by the user requirements. Such requirements can be the area of interest, a time period, the type of the sensor, or the satellite product family.

Their implementation has been successfully applied to two regions (Bolivia and Democratic Republic of Congo) for environmental monitoring.

Finally, Temenos et al. [35] propose a context-aware adaptive data cube (C2A-DC) framework, which utilize Earth Observation data for environmental monitoring and mitigating the climate change effects. Their proposed methodology combines the data cubes formation with calculation of remote sensing operations, including deep and machine learning algorithms for data classification and harmonization. Thus, their outcome is a context-aware adaptive framework, which successfully allows the environmental monitoring and climate change effects based on user predetermined preferences.

2.1 Contribution

Considering the current literature is is noted that several DC implementations have been proposed for different applications over specific regions. However, these implementations are focusing only on the harmonization and homogenization of data provided from the same source, even if they provide DCs of several data sources. As an example, the current DC implementations can provide different data like Landsat-8, Sentinel-2, Sentinel-3, etc., however each cube contains data only from one source. In addition, the current DC implementations provide additional products (i.e., NDVI, NDWI, etc.) on demand based on some query selection.

This work approach the problem of handling the big data from a different scope. The proposed Ocean-DC implementation provide data harmonization and homogenization in a 4D-DC scenario, including data from several sources with commonalities in the same cube. For example, in the case study of this manuscript are used Landsat-8 and Sentinel-2 data for the time series analysis, that are included in the same 4D-DC. In addition, the Ocean-DC precalculates the most common remote sensing products (e.g., NDVI, NDWI, Land Surface Temperature, etc.) providing to the end-users a ready to use product for addressing several environmental and climate change challenges.

3 PROPOSED OCEAN-DC FRAMEWORK

The main objective of the Ocean-DC implementation is to provide harmonization and homogenization between the several available remote sensing and climate data products (i.e., Earth Observation [EO] products), that can be retrieved from multiple data providers (i.e., USGS, Copernicus, NASA). In addition, non-EO data can be used, like the boundaries of the area of interest in the processes. All these data are collected in the first step of the whole framework architecture, as illustrated in Figure 2.

To be more precise, the Data Collection considers several EO and non-EO data including geolocated raster remote sensing and climate products and the non-EO products are usually vector type. Currently, the Ocean-DC implementation supports products from USGS [37] (i.e., Landsat-8,9), Copernicus Services [3] (i.e., Sentinel-2,3,5P and CERRA climate data) and NASA [24] (i.e., MODIS) EO data. The supported non-EO data are polygon shapefile (SHP) geometries. These data are fed to the next step (i.e., Data Harmonization/Homogenization) for generating the Ocean-DCs.
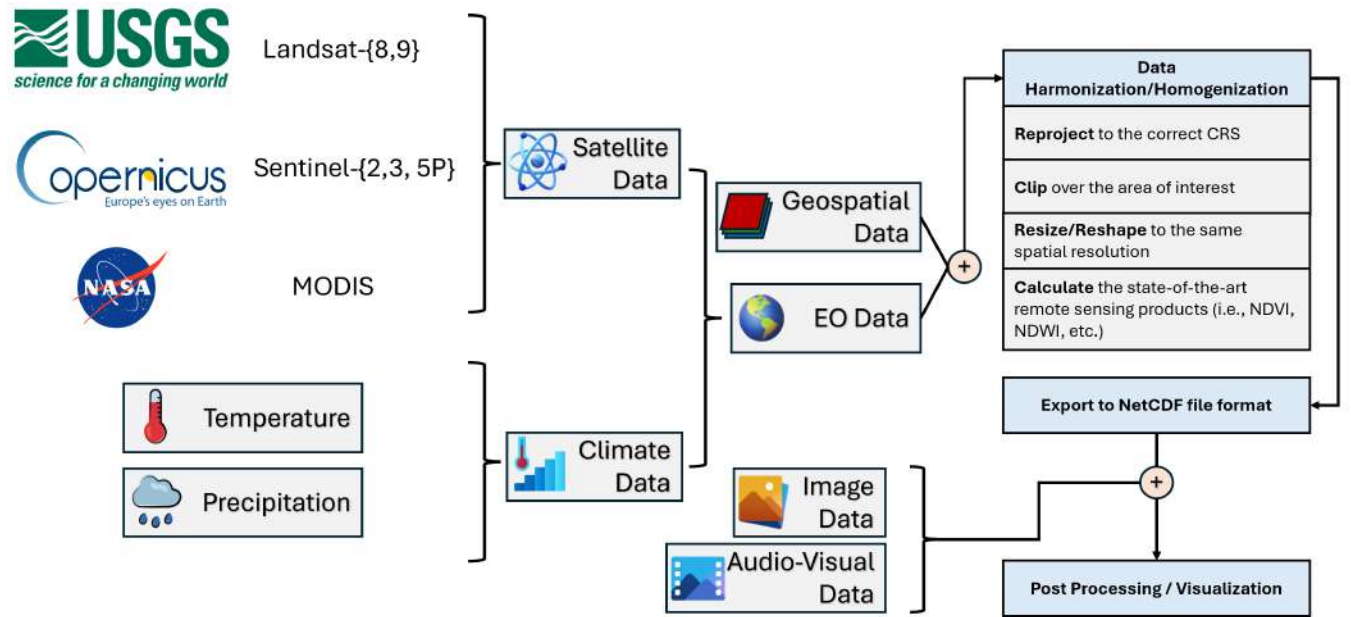
Figure 2: The workflow architecture of the Ocean-DC implementation.

The Data Harmonization/Homogenization step is the core tool [1] of the workflow architecture. In this step, the EO products are reprojected to a common Coordinate Reference System (CRS) to get the same coordinates per pixel. Usually, this coordinate system is the same of the SHP geometry, but the user can select any CRS or EPSG. This step is necessary because the EO data are in different CRS. For instance Landsat-8 and Sentinel-2 raw data are provided in different CRS. Thus, this step is necessary before applying any other process.

After the data reprojection, the framework clips the products to a box geometry based on the user input (i.e., SHP geometry). The box geometry is calculated by the minimum and maximum (x,y) coordinates of this SHP geometry. Image clipping is a critical step, because it significantly minimizes the size of the original images and as a natural result the size of the final Ocean-DC product. After successfully clipping the raster EO products, they are in the same coordinate reference system and visualize the area of interest.

The next following two steps are the image resizing/resampling and the calculation of the most used remote sensing product, in the case of Landsat-8 and Sentinel-2 EO products. The image resizing/resampling process corrects the spatial resolution of the images and ensures that all images have the same width and height pixels. Note that it is not added new information to the original data. This means that an image with spatial resolution 60m/px will be resampled to 10m/px by splitting each pixel to a 6x6 grid with the same value (i.e., pseudo-resolution). This step increases the size of the final product, but it is very important, because the data cube cannot be constructed with different image resolutions. In addition, it is impossible to perform raster calculation with images in different resolutions. Thus, the image resizing/resampling is a mandatory step to proceed in the next steps.

Finally, the calculation of the most common remote sensing products is performed when it is necessary. To be more precise, the remote sensing indices are calculated when specific band information is known (i.e., optical bands, near-infrared, SWIR, etc.). Currently, this step is applied for Landsat-8,9 and Sentinel-2 products. Table 1 summarizes the satellite calculated products, which are containing in an Ocean-DC implementation. The processing of the other products is terminated in the image resizing/resampling step.

The implementation of the Ocean-DC is performed using Python and the state-of-the-art and open-source DC handling library Xarray [39]. This library is empowered with the ready to use methods of rasterio [2], which is another python library for handling EO data. The final outcome of the data harmonization/homogenization framework is a NetCDF file, which is a state-of-the-art format for geodata and is supported by several GIS software like QGIS [7].

Finally, the Ocean-DC products can be combined with other non-EO data like images or drone footage to produce advanced cartographic products or monitor several phenomena. Some examples of post-processing applications can be the monitoring of a disaster (i.e., flooding, forest fire, oil-spill, etc.) or the generation of thematic maps that describes a phenomenon. Other application could be the monitoring of air quality or climate factors (i.e., temperature, precipitation).

4 CASE STUDY ANALYSIS OF THE OIL-SPILL IN SARONIKOS GULF AT SEPTEMPBER 2017

For the demonstration of the Ocean-DC framework is used the oil-spill disaster that took place in Saronic gulf in 10th of September 2017. This disaster has been also investigated in the work of Kolokoussis and Karathanassi [16] where they use the NDWI among

Table 1: Summary of the products, where containing in an Ocean-DC file.

BAND ID	Description	BAND ID	Description
1	COASTAL AEROSOL	23	LSWI-1
2	BLUE	24	LSWI-2
3	GREEN	25	ARVI
4	RED	26	MSAVI2
5	NIR	27	MTVI2
6	SWIR-1	28	VARI
7	SWIR-2	29	TGI
8	CIRRUS	30	LST-1
9	VRE-1	31	LST-2
10	VRE-2	32	LST-CELSIUS
11	VRE-3	33	LST-FHRENHEIT
12	VRE-4	34	VCI
13	WATER VAROUR	35	MNDWI-1
14	PANCHROMATICS	36	MNDW-2
15	TIRS-1	37	WRI-1
16	TIRS-2	38	WRI-2
17	NDVI	39	NDTI
18	NDWI	40	AWEI
19	NDVI (VRE-1)	41	OSI
20	NDVI (VRE-2)	42	NBR-1
21	NDVI (VRE-3)	43	NBR-2
22	NDVI (VRE-4)		

other indices to visually enhance the critical areas of the oil-spill. In a similar way, this work investigates the oil-spill disaster emphasizing mostly on the possibilities that Ocean-DC framework provides to the end-users. For this reason, in this analysis it is used a dataset composed of Landsat-8 and Sentinel-2 products in a time-period before the oil-spill (2016 to 9th of September 2017), some days and weeks after the the oil-spill (10th of September to October 2017) and recent images (2023-2024). Thus, the implemented Ocean-DC product is consisted of 16 satellite imagery products (5 Landsat-8 and 11 Sentinel-2). For visualizing the oil-spill effectively it used the NDWI, WRI-2, and OSI [40] indices for monitoring the water quality of Saronic gulf during the aforementioned periods.

The results of the generated Ocean-DC product are illustrated in Figure 3. The visualization of the generated product performed inside QGIS software and considers the natural (RGB) color composition, the NDWI, WRI, and OSI indices illustrated using pseudo-palette colors. The NDWI product is calculated using the Equation 1 and the output's image values are in range $[-1.0, 1.0]$ as follows: $[-1.0, -0.3) \rightarrow$ Drought, Non-Aqueous Surfaces; $[-0.3, 0.0) \rightarrow$ Moderate drought, non-aqueous surfaces; $[0.0, 0.2) \rightarrow$ Flooding, humidity; $[0.2, 1.0] \rightarrow$ Water Surface [25]. In most of cases, NDWI values in range $[0.0, 0.2)$, indicating the humidity over the Saronicos gulf. Near the coastal areas the values are higher. The only exception is observed in the image of 13th of September, 2017, where it is observed the abnormality in of the oil-spill in the middle of the Saronic Gulf.

$$NDWI = \frac{NIR - SWIR - 1}{NIR + SWIR - 1} \qquad (1)$$

The WRI index is calculated by the Equation 2 and indicates the water ratio in range $[-1.0, 1.0]$ as follows: $[-1.0, -0.75) \rightarrow$ Critical water areas; $[-0.75, -0.25) \rightarrow$ Normal water quality $[-0.25, 0.00) \rightarrow$ Wet ground / Vegetation; $[-0.25, 0.00) \rightarrow$ Wet ground / Vegetation; $[0.00, 1.00] \rightarrow$ Ground / Infrastructures. Based on the observations of the WRI in Figure 3 is is clearly observed a periodic phenomenon of critical water areas among the Salamina coast line. Moreover, in the image of 13th of September, 2017 critical values are observed in the area of the oil-spill (green box in Figure 3), as well as similar areas are indicated in the following weeks after the oil-spill. Another observation of the analysis, is that Landsat-8 products were unable to indicate the oil-spill due to low spatial resolution (i.e., 30m/px).

$$WRI - 2 = \frac{GREEN - RED}{NIR + SWIR - 2} \qquad (2)$$

The OSI (Oil-Spill Index) is referred in the work of [40] and is calculated by the Equation 3. A first observation is the difference of the products' values between Landsat-8,9 and Sentinel-2 data. In Landsat-8,9 products the index effectively distinguish the area of interest into water bodies ($values > 2.5$), vegetation regions ($1.9 < values < 2.5$), and building/infrastructures ($values < 1.9$). In the Sentinel-2 products can be divided to the following ranges: $[0.75, 1.00] \rightarrow$ vegetation regions; $[1.00, 1.9) \rightarrow$ building/infrastructures; and $values > 1.9 \rightarrow$ water bodies. In the water bodies are additionally observed two trends: (a) $[1.9, 2.5) \rightarrow$ indicate areas emitting strongly in blue spectral radiance; (b) $> 2.5 \rightarrow$ indicate areas where water absorbs fully the blue spectral radiance. Considering the aforementioned observations in the product corresponding to the 13th of September, 2017, oil-spill patterns are indicated.

$$OSI = \frac{RED + GREEN}{BLUE} \qquad (3)$$

Combining all the information that is observed by all these indexes, the critical area is bounding by the green box as illustrated in Figure 3. In addition, it is necessary to mention that this time-series analysis indicates other information as well. For instance, among Salamina's coast line is is observed a periodic and recursive pattern of dark pixel values in NDWI and WDI products. This phenomenon indicates additional areas of interest that may be useful for specific use cases related to climate change. Moreover, due to the repeatability of this phenomenon, it is safely to assume that is not correlated with the oil-spill disaster, because it is also observed in images before the oil-spill.

5 DISCUSSION

The main objective of this work is to introduce a new Data Cube implementation for coastal region monitoring named Ocean-DC. The Ocean-DC framework handles several kinds of data combining both EO and non-EO products for generating a NetCDF file that contains all the necessary information over a predetermined area of interest. In addition, during the cube's generation it calculates the most common remote sensing products to help more the end-users by providing them with a ready to use product. As an example, in this paper it was generated a 4D Ocean-DC for monitoring the oil-spill disaster that took place at 10th of September, 2017. The generated Ocean-DC product successfully applied data from different

I. Kavouras, et al.
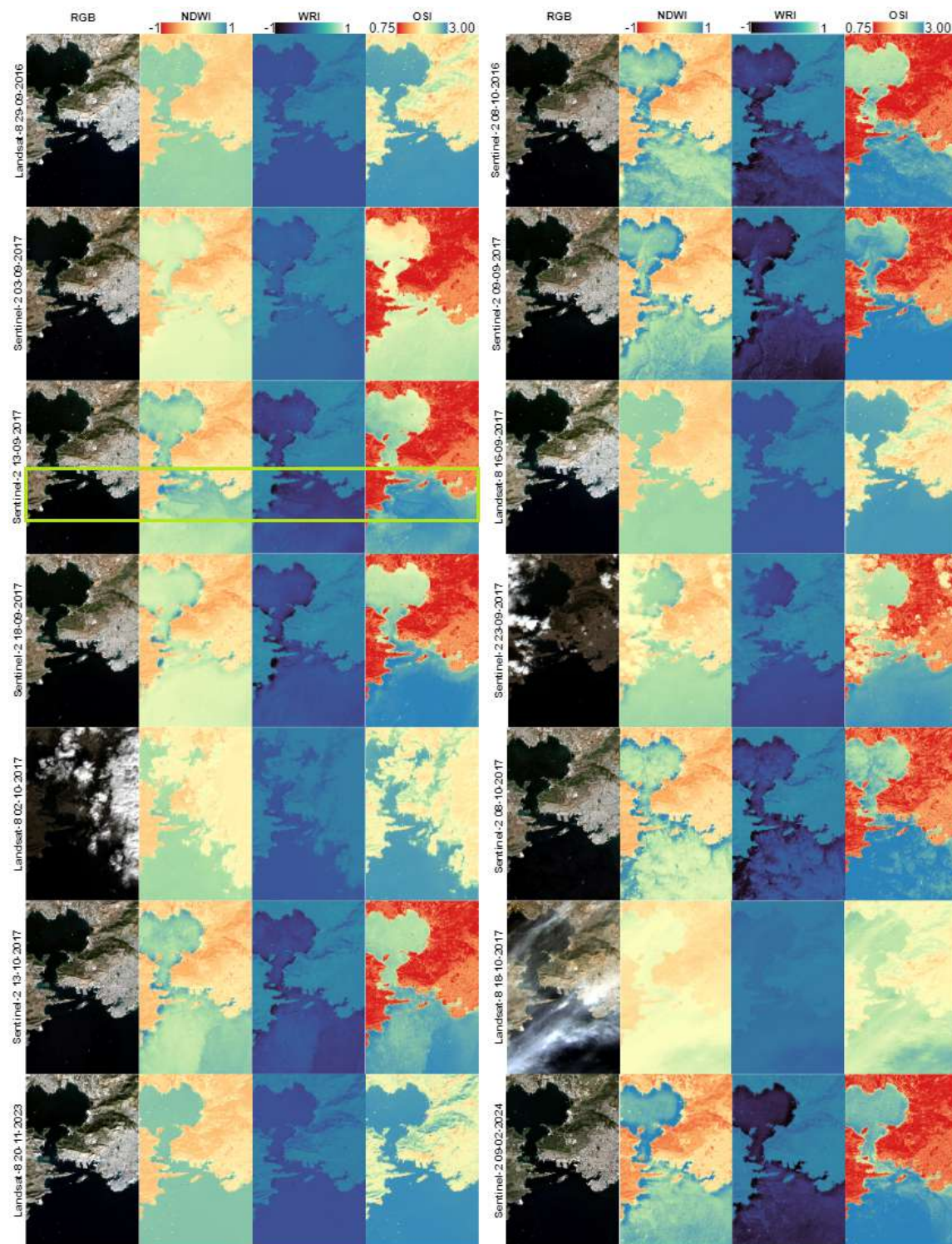
Figure 3: Time series monitoring of the water quality of Saronikos gulf before and after the oil-spill using the RGB, NDWI, WRI, and OSI products of the Ocean-DC implementation file. The green box indicates the area of the oil-spill.

sources (i.e., Landsat-8,9 and Sentinel-2) to a common template and provided a strong ready to use dataset for time series analysis. Thus, the advantages of the Ocean-DC framework are the following:

(1) It utilizes several EO and non-EO data for generating a ready to use product.
(2) It supports several EO data from different sources and harmonize/homogenize them to a common. This feature significantly accelerates the post-processing analysis by allowing the end-users to visualize several data using a single file.
(3) It calculates the state-of-the-art remote sensing data providing a ready to use product. This feature also accelerates the post-processing analysis by allowing the end-users to emphasize only on the interpretation of the generated products without spending time to calculate the products.
(4) The aforementioned characteristics allowing the usage of the Ocean-DC products by both experts and non-experts, because after the generation of the DC, the end-user have to visualize and interpreter the data.
(5) The outcome of the Ocean-DC framework is a NetCDF file, which is readable by the most common GIS software. Thus, it is supported by both free and open-source and commercial GIS software, allowing to end-users to use their preferred software.

Summarizing the discussion, it is worth mentioning that the Ocean-DC framework is not finalized yet. This means, that the framework is updated frequently by expanding this framework with additional features. Moreover, further testing is necessary for indicating further the efficiency of the Ocean-DC framework. Thus, possible future works will combine the Ocean-DC framework with Machine Learning and Deep Learning algorithms for generating more products like cloud and shadow masks, expanding more the applicability of the Ocean-DC products for coastal monitoring. Moreover, other application like in-land vegetation, deforestation, forest fires, or flooding monitoring will be examined as well.

6 CONCLUSIONS

In conclusion, this work introduces a new DC framework that handles several EO and non-EO data from different sources for periodic monitoring the environmental hazards and climate change effects over the coastal areas. The framework may still be under development, however, it can be effectively utilized for time-series analysis and the visualization of disasters over the coastal areas, as indicated in the presented case study. As a result, the Ocean-DC framework can effectively utilized in monitoring phenomena applications. However, the Ocean-DC as an emerging and currently in development implementation will be updated frequently to support more applications by utilizing novel and smart technologies.

ACKNOWLEDGMENTS

REFERENCES

[1] 2024. Ocean-DC framework. https://gitlab.com/JohnCrabs/oceanides_data_cubes (Accessed on 29 Feb 2024).
[2] 2024. Rasterio Documentation. https://rasterio.readthedocs.io/en/stable/ (Accessed on 29 Feb 2024).
[3] European Space Agency. 2024. Copernicus Hub. https://dataspace.copernicus.eu/ (Accessed on 29 Feb 2024).
[4] Peter Baumann, Dimitar Misev, Vlad Merticariu, and Bang Pham Huu. 2019. Datacubes: Towards space/time analysis-ready data. *Service-Oriented Mapping: Changing Paradigm in Map Production and Geoinformation Management* (2019), 269–299.
[5] Kamal Boulil, Sandro Bimonte, and Francois Pinet. 2015. Conceptual model for spatial data cubes: A UML profile and its automatic implementation. *Computer Standards & Interfaces* 38 (2015), 113–132.
[6] Bruno Chatenoux, Jean-Philippe Richard, David Small, Claudia Roeoesli, Vladimir Wingate, Charlotte Poussin, Denisa Rodila, Pascal Peduzzi, Charlotte Steinmeier, Christian Ginzler, et al. 2021. The Swiss data cube, analysis ready data archive using earth observations of Switzerland. *Scientific data* 8, 1 (2021), 295.
[7] QGIS Community. 2024. Quantum GIS. https://qgis.org/en/site/ (Accessed on 29 Feb 2024).
[8] Karine R. Ferreira, Gilberto R. Queiroz, Lubia Vinhas, Rennan F. B. Marujo, Rolf E. O. Simoes, Michelle C. A. Picoli, Gilberto Camara, Ricardo Cartaxo, Vitor C. F. Gomes, Lorena A. Santos, Alber H. Sanchez, Jeferson S. Arcanjo, José Guilherme Fronza, Carlos Alberto Noronha, Raphael W. Costa, Matheus C. Zaglia, Fabiana Zioti, Thales S. Korting, Anderson R. Soares, Michel E. D. Chaves, and Leila M. G. Fonseca. 2020. Earth Observation Data Cubes for Brazil: Requirements, Methodology and Products. *Remote Sensing* 12, 24 (2020). https://doi.org/10.3390/rs12244033
[9] Matteo Francia, Patrick Marcel, Verónika Peralta, and Stefano Rizzi. 2022. Enhancing cubes with models to describe multidimensional data. *Information Systems Frontiers* 24, 1 (2022), 31–48.
[10] Gregory Giuliani, Bruno Chatenoux, Antonio Benvenuti, Pierre Lacroix, Mattia Santoro, and Paolo Mazzetti. 2020. Monitoring land degradation at national level using satellite Earth Observation time-series data to support SDG15–exploring the potential of data cube. *Big Earth Data* 4, 1 (2020), 3–22.
[11] Gregory Giuliani, Bruno Chatenoux, Thomas Piller, Frédéric Moser, and Pierre Lacroix. 2020. Data Cube on Demand (DCoD): Generating an earth observation Data Cube anywhere in the world. *International Journal of Applied Earth Observation and Geoinformation* 87 (2020), 102035.
[12] Vitor CF Gomes, Gilberto R Queiroz, and Karine R Ferreira. 2020. An overview of platforms for big earth observation data management and analysis. *Remote Sensing* 12, 8 (2020), 1253.
[13] Jim Gray, Surajit Chaudhuri, Adam Bosworth, Andrew Layman, Don Reichart, Murali Venkatrao, Frank Pellow, and Hamid Pirahesh. 1997. Data cube: A relational aggregation operator generalizing group-by, cross-tab, and sub-totals. *Data mining and knowledge discovery* 1 (1997), 29–53.
[14] Jiawei Han, Yixin Chen, Guozhu Dong, Jian Pei, Benjamin W Wah, Jianyong Wang, and Y Dora Cai. 2005. Stream cube: An architecture for multi-dimensional analysis of data streams. *Distributed and Parallel Databases* 18 (2005), 173–197.
[15] Venky Harinarayan, Anand Rajaraman, and Jeffrey D Ullman. 1996. Implementing data cubes efficiently. *Acm Sigmod Record* 25, 2 (1996), 205–216.
[16] Polychronis Kolokoussis and Vassilia Karathanassi. 2018. Oil Spill Detection and Mapping Using Sentinel 2 Imagery. *Journal of Marine Science and Engineering* 6, 1 (2018). https://doi.org/10.3390/jmse6010004
[17] Steve Kopp, Peter Becker, Abhijit Doshi, Dawn J Wright, Kaixi Zhang, and Hong Xu. 2019. Achieving the full vision of earth observation data cubes. *Data* 4, 3 (2019), 94.
[18] Walter Leal Filho, Abdul-Lateef Balogun, Olawale Emmanuel Olayide, Ulisses M Azeiteiro, Desalegn Y Ayal, Pastor David Chavez Muñoz, Gustavo J Nagy, Paulette Bynoe, Otienoh Oguge, N Yannick Toamukum, et al. 2019. Assessing the impacts of climate change in cities and their adaptive capacity: Towards transformative approaches to climate change adaptation and poverty reduction in urban areas in a set of developing countries. *Science of the Total Environment* 692 (2019), 1175–1190.
[19] Julien Leprince, Clayton Miller, and Wim Zeiler. 2021. Data mining cubes for buildings, a generic framework for multidimensional analytics of building performance data. *Energy and Buildings* 248 (2021), 111195.
[20] Adam Lewis, Simon Oliver, Leo Lymburner, Ben Evans, Lesley Wyborn, Norman Mueller, Gregory Raevksi, Jeremy Hooke, Rob Woodcock, Joshua Sixsmith, Wenjun Wu, Peter Tan, Fuqin Li, Brian Killough, Stuart Minchin, Dale Roberts, Damien Ayers, Biswajit Bala, John Dwyer, Arnold Dekker, Trevor Dhu, Andrew Hicks, Alex Ip, Matt Purss, Clare Richards, Stephen Sagar, Claire Trenham, Peter Wang, and Lan-Wei Wang. 2017. The Australian Geoscience Data Cube — Foundations and lessons learned. *Remote Sensing of Environment* 202 (2017), 276–292. https://doi.org/10.1016/j.rse.2017.03.015
[21] Arfan Majeed, Yingfeng Zhang, Shan Ren, Jingxiang Lv, Tao Peng, Saad Waqar, and Enhuai Yin. 2021. A big data-driven framework for sustainable and smart

additive manufacturing. *Robotics and Computer-Integrated Manufacturing* 67 (2021), 102026.

[22] Zhifu Mi, Dabo Guan, Zhu Liu, Jingru Liu, Vincent Viguié, Neil Fromer, and Yutao Wang. 2019. Cities: The core of climate change mitigation. *Journal of Cleaner Production* 207 (2019), 582–589.

[23] Fabio Miranda, Lauro Lins, James T Klosowski, and Claudio T Silva. 2017. Topkube: A rank-aware data cube for real-time exploration of spatiotemporal data. *IEEE Transactions on visualization and computer graphics* 24, 3 (2017), 1394–1407.

[24] NASA. 2024. MODIS Data Hub. https://modis.gsfc.nasa.gov/data/ (Accessed on 29 Feb 2024).

[25] Riyaz Khan NH. 2022. Study of fluctuations in surface area of Lake Haramaya using NDWI and MNDWI methods. *JGISE: Journal of Geospatial Information Science and Engineering* 5, 1 (2022), 36–41.

[26] AG Olabi and Mohammad Ali Abdelkareem. 2022. Renewable energy and climate change. *Renewable and Sustainable Energy Reviews* 158 (2022), 112111.

[27] David E Parker. 2010. Urban heat island effects on estimates of observed climate change. *Wiley Interdisciplinary Reviews: Climate Change* 1, 1 (2010), 123–133.

[28] Sahar Hadi Pour, Ahmad Khairi Abd Wahab, Shamsuddin Shahid, Md Asaduzzaman, and Ashraf Dewan. 2020. Low impact development techniques to mitigate the impacts of climate-change-induced urban floods: Current trends, issues and challenges. *Sustainable Cities and Society* 62 (2020), 102373.

[29] Charlotte Poussin, Yaniss Guigoz, Elisa Palazzi, Silvia Terzago, Bruno Chatenoux, and Gregory Giuliani. 2019. Snow Cover Evolution in the Gran Paradiso National Park, Italian Alps, Using the Earth Observation Data Cube. *Data* 4, 4 (2019). https://doi.org/10.3390/data4040138

[30] Khaiwal Ravindra, Preety Rattan, Suman Mor, and Ashutosh Nath Aggarwal. 2019. Generalized additive models: Building evidence of air pollution, climate change and human health. *Environment international* 132 (2019), 104987.

[31] Sunita Sarawagi, Rakesh Agrawal, and Nimrod Megiddo. 1998. Discovery-driven exploration of OLAP data cubes. In *Advances in Database Technology—EDBT'98: 6th International Conference on Extending Database Technology Valencia, Spain, March 23–27, 1998 Proceedings 6.* Springer, 168–182.

[32] Maocai Shen, Wei Huang, Ming Chen, Biao Song, Guangming Zeng, and Yaxin Zhang. 2020. (Micro) plastic crisis: un-ignorable contribution to global greenhouse gas emissions and climate change. *Journal of Cleaner Production* 254 (2020), 120138.

[33] Ying Song, Yingling Fan, Xin Li, and Yanjie Ji. 2018. Multidimensional visualization of transit smartcard data using space–time plots and data cubes. *Transportation* 45 (2018), 311–333.

[34] Roberto Tardío, Alejandro Maté, and Juan Trujillo. 2020. A new big data benchmark for OLAP cube design using data pre-aggregation techniques. *Applied Sciences* 10, 23 (2020), 8674.

[35] Anastasios Temenos, Nikos Temenos, Ioannis N. Tzortzis, Ioannis Rallis, Anastasios Doulamis, and Nikolaos Doulamis. 2024. C2A-DC: A context-aware adaptive data cube framework for environmental monitoring and climate change crisis management. *Remote Sensing Applications: Society and Environment* 34 (2024), 101171. https://doi.org/10.1016/j.rsase.2024.101171

[36] Anastasios Temenos, Ioannis N Tzortzis, Maria Kaselimi, Ioannis Rallis, Anastasios Doulamis, and Nikolaos Doulamis. 2022. Novel insights in spatial epidemiology utilizing explainable AI (XAI) and remote sensing. *Remote Sensing* 14, 13 (2022), 3074.

[37] USGS. 2024. Earth Explorer. https://earthexplorer.usgs.gov/ (Accessed on 29 Feb 2024).

[38] Hua Wu and Zhao-Liang Li. 2009. Scale issues in remote sensing: A review on analysis, processing and modeling. *Sensors* 9, 3 (2009), 1768–1793.

[39] Xarray. 2024. Xarray Documentation. https://docs.xarray.dev/en/stable/ (Accessed on 29 Feb 2024).

[40] Mohamed Zakzouk, Islam Abou El-Magd, Elham M Ali, Abdulaziz M Abdulaziz, Amjad Rehman, and Tanzila Saba. 2024. Novel oil spill indices for sentinel-2 imagery: A case study of natural seepage in Qaruh Island, Kuwait. *MethodsX* 12 (2024), 102520. https://doi.org/10.1016/j.mex.2023.102520

[41] RenHua Zhang, Jing Tian, ZhaoLiang Li, HongBo Su, ShaoHui Chen, and XinZhai Tang. 2010. Principles and methods for the validation of quantitative remote sensing products. *Science China Earth Sciences* 53 (2010), 741–751.