Remote Sensing Applications: Society and Environment 34 (2024) 101171

Contents lists available at ScienceDirect

Remote Sensing Applications: Society and Environment

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

C2A-DC: A context-aware adaptive data cube framework for environmental monitoring and climate change crisis management ☆

Anastasios Temenos [a,b,*], Nikos Temenos [b], Ioannis N. Tzortzis [a,b], Ioannis Rallis [a,b], Anastasios Doulamis [a], Nikolaos Doulamis [a]

[a] Department of Rural, Surveying and Geoinformatics Engineering, National Technical University of Athens, Iroon Polytechneiou 9, Athens, 157 73, Greece

[b] Institute of Communications and Computer Systems, Iroon Polytechneiou 9, Athens, 157 73, Greece

ARTICLE INFO

Keywords:
Data cubes
Analysis-ready data
Earth observation
Remote sensing
AI
DL
ML
Cloud detection
Deforestation

ABSTRACT

A context-aware adaptive data cube (*C2A-DC*) framework based on Earth Observation (EO) data for environmental monitoring to mitigate Climate Change (CC) effects is proposed. It has the property of combining DC formation, calculation of Remote Sensing (RS) operations, deep and machine learning algorithms for classification and data harmonization by updating the layers of the DC using information obtained from the previous operations, all applied to EO data originating from a satellite selected by a stakeholder. Moreover, the proposed framework is context-aware adaptive in the sense that it allows for environmental monitoring tasks according to a stakeholder's needs, which is in contrast to existing works being constrained to a single type of environmental monitoring. Furthermore, with the proposed framework end-users do not have to make any additional processing action to handle the data as they are given to them as a set of analysis-ready data, harmonically aligned under a unique DC. To showcase the *C2A-DC* framework's effectiveness and applicability in environmental monitoring, crisis management, and post-disaster assessment, it is applied to two real-world cases. The first one corresponds to cloud and shadow detection in a flood event, with several Convolutional Neural Network (CNN) architectures being trained on the C2S-MS floods dataset for classification, while k-means clustering and RS product calculations are used for monitoring scenery changes. The second one corresponds to deforestation monitoring from fires, evaluated using RS products.

1. Introduction

Modern urban plans are designed with the purpose of interconnecting people, infrastructures, activities and many other resources into close-distance building blocks, making them susceptible to Climate Change (CC) effects (Mi et al., 2019; Leal Filho et al., 2019; Olabi and Abdelkareem, 2022). Daily human activities are linked with products and services that derive from burning fossil fuels, forest cutting, farming livestock and many more, all having a negative impact on the climate (Shivanna, 2022; Reza and Sabau, 2022). All these activities result in (a) lack of urban greenness (Temenos et al., 2022), (b) urban flooding (Pour et al., 2020), (c)

☆ This work was financially supported by the Horizon 2020 European Union project HARMONIA "Development of a Support System for Improved Resilience and Sustainable Urban areas to cope with Climate Change and Extreme Events based on GEOSS and Advanced Modelling Tools" funded under the call H2020-LC-CLA-19-2020, with grant agreement No. 101003517.

* Corresponding author at: Department of Rural, Surveying and Geoinformatics Engineering, National Technical University of Athens, Iroon Polytechneiou 9, Athens, 157 73, Greece.

E-mail addresses: tasostemenos@mail.ntua.gr (A. Temenos), ntemenos@gmail.com, niktemenos@mail.ntua.gr (N. Temenos), itzortzis@mail.ntua.gr (I.N. Tzortzis), irallis@mail.ntua.gr (I. Rallis), adoulam@cs.ntua.gr (A. Doulamis), ndoulam@cs.ntua.gr (N. Doulamis).

https://doi.org/10.1016/j.rsase.2024.101171
Received 19 September 2023; Received in revised form 14 December 2023; Accepted 2 February 2024
Available online 16 February 2024

absence of high Atmospheric Quality (AQ) (Ravindra et al., 2019), (d) increased Green House Gas (GHG) emissions (Shen et al., 2020), (e) geo-hazards such as ground deformation due to soil degradation (Tsatsaris et al., 2021) and (f) heat islands (Parker, 2010). In addition, extreme population shifts from rural areas to urban centers intensify the urbanization phenomenon and CC effects (Adger et al., 2020; Kirikkaleli and Sowah, 2020). The impact of the above crisis depends on how prepared the city is *to respond to potential specific factors against the CC effects*, leaning towards becoming resilient and protective during crises (Bush and Doyon, 2019; Afrin et al., 2021), but also remaining functional.

Towards an efficient design and implementation of a decision-based response system, Remote Sensing (RS) and Artificial Intelligence (AI) are recently utilized. Specifically, the concept of Data Cubes (DC) has been introduced by the research community to provide *ready-to-use time series data*, captured from satellite infrastructures, to environmental experts for assessing the CC effects and, in the sequel, for delivering solutions for hazard and crisis management. DC is a multi-dimensional array-based structured framework of a unique coordinate reference system consisting of *spatio-temporal and spectral information* to enable efficient environmental monitoring. DC enables the exploitation of remote sensing data of a region, appropriately pre-processed according to specific requirements of an application scenario (e.g., flood mapping or forest fire management), to be ready for emerging environmental crisis management and rescue planning during natural hazards (Gorelick et al., 2017; Tamiminia et al., 2020).

The different application scenarios of environmental crisis and natural hazard management (e.g., fire, flooding, and soil degradation) are, however, inherently heterogeneous in their requirements, and consequently, they demand different forms of information/data, prediction models, data descriptors (e.g. features, indices) and workflows. Therefore, different structures, forms, and methods (e.g. AI-driven prediction models) of DC are necessary to address the quite different needs of an application scenario. To tackle the aforementioned difficulties, an alternative context-aware adaptive DC structure, called as *C2A-DC* in the following, is proposed in this paper, enhanced with adaptable capabilities to face the specific requirements of an environmental crisis and natural hazard application scenario. *C2A-DC* is a context-aware data structure in the sense that it incorporates re-configurable forms with the main purpose to adjust the captured data and appropriately scaling and geo-referencing to fit the variety of environmental tasks and post-disaster assessment. The main structure of *C2A-DC* is the following:

a *Acquisition Context*: This is an interoperable XML schema for dynamic information gathering from different satellite images to fit the needs and requirements of an application scenario considered. For example, regarding flooding management SAR-driven information is necessary (apart from visual sensory data), while regarding immediate fire monitoring high temporal frequent data is mandatory, like information provided by MODIS.

b *Abstract and Clustering Context*: The purpose of this layer is to reduce the high dimensionality of the acquired information, through the application of unsupervised (e.g., clustering) and self-supervised methodologies, to provide meaningful ready-to-use information to environmental experts.

c *Descriptor and Processing Context*: This layer refers to the specific processing methods on the captured data to extract key features as ready-to-use information for a particular application scenario.

d *Data Harmonization and Geo-Reference Context*: The acquired data need to be harmonized together so that the environmental experts have the most appropriate spatio-temporal information in a unique coordinate geo-referencing system.

e *AI Adaptable Context*: Re-configurable pool of different machine learning and neural networks models are utilized in the proposed *C2A-DC* structure to allow for a high level semantic processing of the acquired information towards prediction, modeling and spatio-temporal mapping in the context requirements of a specific application scenario.

f *Decision-Making Context*: Fusion technologies are considered to exploit the outcomes of all the aforementioned context-aware layers to conclude into near real-time decision making recommendations for natural hazard crisis management.

1.1. Previous works

One of the first implementations of DC in EO Science was the Australian Geoscience DC (AGDC) (Lewis et al., 2017). AGDC addresses the Big Data challenges including volume, velocity, and variety, limiting the EO data usefulness. AGDC has been successfully applied for several applications including (i) water observations from space, (ii) forest cover change, (iii) coastal erosion, (iv) agriculture monitoring and (v) biodiversity (Lewis et al., 2017).

In Chatenoux et al. (2021), authors present how the Swiss DC (SDC) is built on the Open DC (ODC) software, which is a geospatial data management and analysis open-source software for handling satellite data. The ODC provides a framework to access, store, manage, and analyze grid-aligned satellite EO data collections in large quantities. The SDC uses the ODC to provide easy access and tools to analyze synoptic data, so as to divide them better into categories, e.g. climate, and agriculture to name a few.

Authors in Poussin et al. (2019), exploit the advantages of the EO SDC to monitor snow cover change in Gran Paradiso National Park, Alps. They aim to study the characteristics of the snow dataset, such as satellite observations of cloudless images, as well as stress the strengths and weaknesses of the DC to monitor snow cover change. They utilize remotely sensed data and time series analyses using Landsat satellite observations stored in an ODC. One of the major limitations of the snow detection algorithm used in the study is that optical sensors struggle to obtain surface information when clouds exist. Although using mosaicking image processing techniques and multi-sensor analysis, clouds are still present limiting the number of available observations to map snow covering.

Another work utilizing the SDC is presented in Giuliani et al. (2020a), proposing a methodology to generate accurate and consistent land degradation data using satellite imagery. The proposed approach uses a DC and a set of sub-indicators to assess land productivity and cover, and soil organic carbon changes to evaluate land degradation.

Authors in Giuliani et al. (2020b) propose the DC on Demand (DCoD), a script-based tool to automatically generate a DC under given user requirements such as area of interest, temporal extent, and type of sensor. It has been successfully applied to two sites in Bolivia and the Democratic Republic of Congo (DRC) for environmental monitoring.

Another worth utilizing DC implementation is the Africa Regional DC (Killough, 2019), which is based on the ODC. It is focused on the development of the capacity of end-users to apply analysis-ready EO data for local and national needs. Some of the key application areas that the ODC initiative is addressing in Africa include deforestation, water quality, agriculture and others.

Authors in Ferreira et al. (2020) propose an efficient way to generate EO DCs for the Brazil region exploiting the benefits of the ODC platform. Their contribution is narrowed to the conduction of analysis-ready data (ARD) and multidimensional DCs from RS images to efficiently map land use and land cover by employing image time series analysis and machine learning techniques.

Despite their wide range of applications and efficient EO data management, all the approaches discussed so far using DCs are static; once the DCs are formed from image time-series, to further process information an expert has to acquire the DC and then proceed with post DC formation processing. This means that RS products such as the NDVI and others are not included in the DCs as additional layers. Therefore, the limitations introduced to individuals are twofold; on the one hand, an end-user is limited from assessing environmental tasks directly, while on the other hand, an expert has to put additional effort into the calculation of RS products.

1.2. Contribution

In this work, we propose a new context-aware adaptive Data Cube structure, called *C2A-DC*, as an analysis-ready data arrangement to fit the heterogeneous requirements of different environmental crisis application scenarios. In contrast to the traditional static Data Cubes, *C2A-DCs* includes several context-aware layers to allow the acquisition of different information, based on the needs of the application scenario considered, as well as processing, organization (including geo-referencing), and AI-based decision making. The advantage of the proposed *C2A-DC* is its re-configurable design capability to facilitate its usage over heterogeneous application scenario requirements. In particular, the main contribution of the present work is as follows:

- The introduction of an interoperable XML schema to allow for dynamic data acquisition from different satellite and ground stations sensors to fit the particularities of a specific application scenario.
- Semantic categorization of RS indices with respect to the needs of a natural hazard scenario. For example, green indices are assigned to be suitable for deforestation modeling, while blue indices are for water oriented environmental management.
- Incorporation of advanced abstraction and clustering methodologies, based on the research of unsupervised neural networks to reduce the high-dimension information captured by the acquisition layer to enhance decision making and data processing.
- Incorporation of advanced adaptable neural network structures and a memory of a pool of deep learning modules to allow for a domain-specific semantic, high-level processing of the acquired information.
- Data harmonization toolkits to geo-reference the acquired heterogeneous data into a common a unique coordination system, to ensure temporal spatial monitoring of a region of interest.
- Application of the proposed context-aware adaptive DC (*C2A-DC*) into two real-word cases to illustrate how the proposed framework can be used to tackle CC effects onto heterogeneous application scenario domains.

The rest of this manuscript is organized as follows. Section 2 introduces the proposed *C2A-DC* framework and explains its operation principle, while Section 3 describes in detail the dataset, the CNN architectures and the clustering algorithms as well as the performance evaluation metrics of the models. Section 4 proceeds with the use of the proposed *C2A-DC* framework in two real-world case studies including (1) cloud and shadow detection and (2) deforestation monitoring. Finally, Section 5 concludes the proposed work and discusses potential extensions.

2. Proposed *C2A-DC* framework

The proposed *C2A-DC* framework is shown in Fig. 1. Initially, a stakeholder requests EO data to perform a desired task on an event, e.g. environmental monitoring, post-disaster assessment, etc. The EO data themselves can potentially be heterogeneous, meaning that they might contain information from different sensing element sources such as satellite images, temperature readings, etc., thus making the DC a suitable choice for structuring them properly for processing. As part of the request, the stakeholder specifies the temporal extension of the event and its corresponding spatial boundaries and then the framework forms a DC using the ODC platform. Before the established DC is given back to the expert, it passes through image processing operations, including RS indices creation, as well as Machine and Deep Learning processing, including classification tasks and unsupervised clustering, as illustrated in detail in Fig. 2. The processed outputs are then aligned into the formed DC under the same coordinate system in the selected time interval. As such the framework provides experts with analysis-ready data and enriched layers that enhance environmental monitoring.

The operation principle of the proposed framework in Fig. 1 can be separated into the following stages: (1) data acquisition and data cube formation, (2) image processing operations and AI-based processing, and (3) data harmonization. Each one of the above is explained below.

(1) Data acquisition and data cube formation: A multispectral image $I(m, l)$, where m, l refer to the width and height of the image respectively, consists of $k = 1, \ldots, K$ channel bands $B_k(m, l)$ capturing different wavelengths of the spectrum. Multispectral

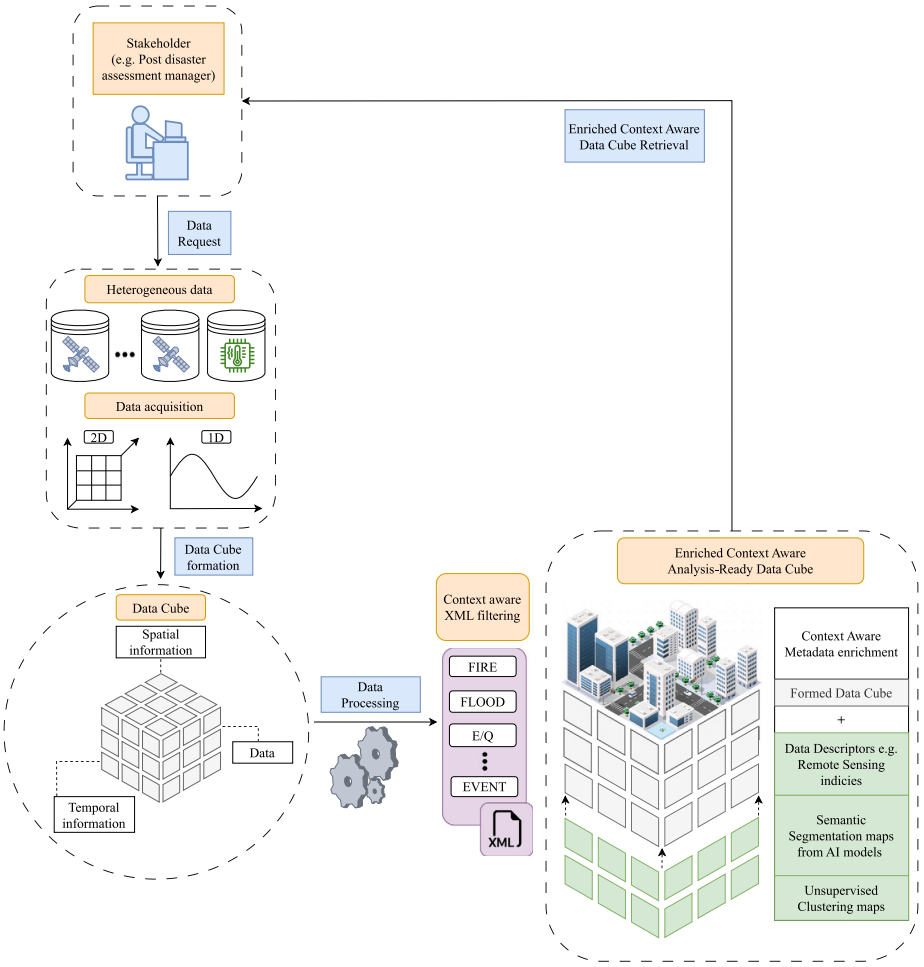

Fig. 1. Workflow of the proposed C2A-DC framework. Upon request by a stakeholder, C2A-DC gets EO heterogeneous data and forms them as a DC to proceed with data processing. Once the processing is finished, the data are harmonized under the same coordinate system and are sent back to the stakeholder as an enriched context-aware adaptive DC.

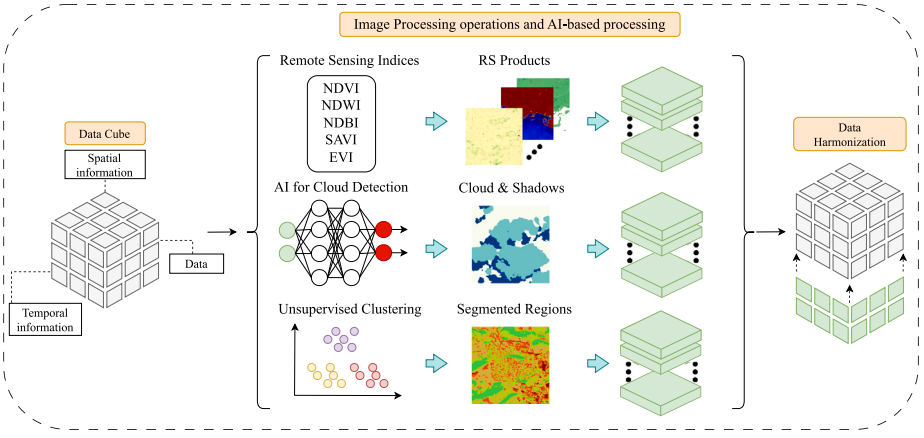

Fig. 2. Data processing operations performed by the proposed C2A-DC of Fig. 1, including Image Processing operations, RS indices creation, Machine and Deep Learning for classification and unsupervised clustering tasks. The output results are aligned with the initial layers of the DC during data harmonization.

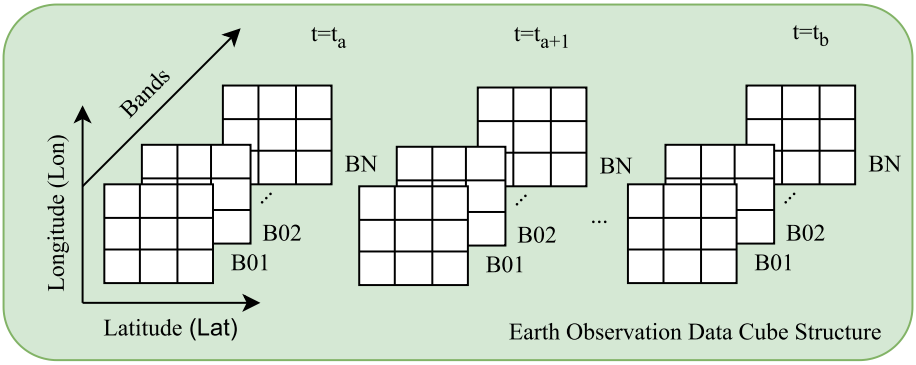

Fig. 3. Demonstration of the $k = 1, \dots, K$ bands existing an EO DC with their pixel values being updated during a selected time frame $t = t_a, \dots, t_b$.

images from satellites contain in each pixel beyond the spectral signature information regarding the latitude x and longitude y of the geospatial extension under a unique coordinate system. According to the satellite's orbit, such images are captured in two arbitrarily selected time stamps $t = t_a, \dots, t_b$, with $t_a < t_b$ with the purpose of projecting the same land scenery with coordinates (x, y). An EO DC is defined as a unified structure formed by aligning these images sequentially given as

$$\mathrm{DC} = \begin{bmatrix} I_{(x,y)}(m_t, l_t, t = t_a) \\ I_{(x,y)}(m_t, l_t, t = t_a + 1) \\ \vdots \\ I_{(x,y)}(m_t, l_t, t = t_b - 1) \\ I_{(x,y)}(m_t, l_t, t = t_b) \end{bmatrix}, \tag{1}$$

where t_a, t_b represent the initial and the final selected time stamps respectively, while $I_{(x,y)}(m_t, l_t, t)$ represents an image referring to a specific land scenery (x, y) whose pixel values change over time t. A graphical illustration of the above definitions is shown in Fig. 3.

(2) Image Processing operations and AI-based processing: Once the data are collected and the DC is formed, it can be used for further processing using image processing operations focusing on the RS indices, ML and DL algorithms. The importance of the DC in each category is explained below.

With respect to the RS indices that enhance the information of basic land categories, namely water, vegetation, urban fabric and soil, land cover change and its monitoring can be accelerated by utilizing them. The RS indices usually include the normalized difference vegetation index (NDVI), the normalized difference water index (NDWI), the normalized difference build-up index (NDBI), the soil-adjusted vegetation index (SAVI), the enhanced vegetation index (EVI) and others. Using the DC allows for direct computation of these RS indices for a selected time frame while maintaining the georeference information (x, y), which is in contrast to the traditional desktop GIS applications requiring manual procedures for their computation and storing.

Maintaining the georeference information comes also in handy in the case of ML algorithms, with emphasis given on clustering; using k-means clustering on a land scenery enables the segmentation of its land use over a selected time frame in an unsupervised manner considering only its spectral signature. For instance, one can identify how vegetation and soil are segmented and in how many subcategories they can be further classified such as crops and forests. K-means is an easy-to-implement algorithm that scales to large datasets, including remote sensing ones where multiple channels are considered. Its major challenge is the initialization of K-classes, which is addressed in the proposed framework by creating maps and increasing linearly the number of classes k.

Despite the data from the DC being ready for further processing, they can in fact contain information that limits the visibility of the land scenery, which is the case when clouds and their shadows are present. To that end, DL can be used to tackle such issues, with Convolutional Neural Networks (CNNs) being effective in segmenting the clouds and their shadows from the rest land scenery. As such, the expert is able to discard images with increased cloud coverage so as to utilize information from other ones in the selected time window.

(3) Data harmonization: One of the strengths of the proposed *C2A-DC* framework lies in its data harmonization stage. In particular, the information obtained from the RS operations, the clustered maps segmenting different land uses and the clouds with their shadows being detected and localized, is returned and saved within the DC in the form of new layers on top of the existing ones. In other words, assuming that k' new channels $C_{(x,y)}(m, l, t')$ are computed and considering the initial DC information from (1), the

enriched DC is described as

$$\mathrm{DC} = \begin{bmatrix} I_{(x,y)}(m_t, l_t, t = t_a) \\ I_{(x,y)}(m_t, l_t, t = t_a + 1) \\ \vdots \\ I_{(x,y)}(m_t, l_t, t = t_b - 1) \\ I_{(x,y)}(m_t, l_t, t = t_b) \\ C^1_{(x,y)}(m, l, t' = t'_a) \\ C^2_{(x,y)}(m, l, t' = t'_a + 1) \\ \vdots \\ C^{k'}_{(x,y)}(m, l, t' = t'_b) \end{bmatrix}. \tag{2}$$

Note that t' is used to denote the time-index of an arbitrarily selected time frame within $t_a, \ldots, t_b$, which is such that $t_a \leq t'_a < t'_a + 1, \ldots t'_b \leq t_b$.

Considering the operation of the proposed *C2A-DC* framework, it can be concluded that it benefits both experts and end users. On the one hand, the expert can either create or use it for initial urban planning purposes by selecting a time period, adding and/or modifying data existing within the DC, as well as monitoring events that may occur and later saving them to cloud for further use. On the other hand, any end user can download information contained in the DC, without following the typical procedures of downloading, storing and processing data locally.

3. Experimental setup

In this section we proceed with the description of the dataset, the AI models and the classification metrics used.

3.1. Dataset description

The performance of the proposed *C2A-DC* framework is evaluated using the C2S-MS Floods[1] dataset, containing annotated global flood events taken from the Sentinel-2 satellite. Within the dataset, there are 900 image patches having 512×512 pixel size, covering $K = 13$ spectral bands. The image patches are selected from 18 global flooding events, containing annotations that target water bodies, clouds, and their shadows.

The motivation behind the dataset selection lies within the natural disaster phenomena that it illustrates. Specifically, water bodies are confused with clouds and their shadows due to similarities between their spectral signatures (Li et al., 2021), leading to miss-classifications in flood mapping when rule-based models are used for such purposes.

We consider two experiments in this work, both utilizing three band combinations. In the first one, we consider B8 A (Near Infrared) - B11 (SWIR1) - B12 (SWIR2), all having 20-m resolution per pixel, while in the second experiment, we consider the Color Infrared band combination B08 (Near Infrared) - B04 (Red) - B03 (Green), all having 10-m resolution per pixel. By doing so, the effect of DL architectures on the increase of the resolution in meters per pixel can be investigated.

One can argue that using only 900 image patches can be a small number for the generalization of the models' predictability. To this end, we augment the initial dataset by splitting each patch into 4 sub-regions recreating from each image 4 new ones with halved size e.g. from 512×512 to 256×256. Moreover, we consider both horizontal and vertical flip. The initial dataset with parameters (Number of images, image width, image height, channels) is therefore transformed from $(900, 512, 512, 13)$ to $(10788, 256, 256, 3)$.

3.2. AI model parameters

Deep Learning Methods: Here we discuss the architecture of each one of the DL models we have used in the experiments of our proposed *C2A-DC* framework. These include: (A) UNet (Ronneberger et al., 2015), (B) R2UNet (Alom et al., 2018), (C) Attention UNet (Oktay et al., 2018), (D) Attention R2UNet (Zuo et al., 2021), (E) DeepLabV3+ (Chen et al., 2018) and (F) LinkNet (Chaurasia and Culurciello, 2017). Moreover, we summarize their parameters used in Table 1.

(A) UNet: It is a fully CNN consisting of a contracting and an expanding path. The contracting path is used for down-sampling and has five convolution (Conv) layers with double 3×3 convolutions, followed by a rectified linear unit (ReLU) activation function and a 2×2 max pooling operation of stride 2. The expanding path used for up-sampling has the same number of Conv layers as the contracting path, but, includes concatenation of the feature maps stemming from the contracting path (Ronneberger et al., 2015). The final layer of the expanding path uses a 1×1 convolution to map the final feature vector to the desired number of classes. Note here that the base number of filters used at each Conv layer is 16 and is doubled after each step.

(B) R2UNet: It follows a similar structure as that of the UNet, consisting of 5 contracting and expanding paths. Compared to UNet, R2UNet utilizes recurrent residual Conv layers; inputs are sequentially fed to two recurrent Conv + ReLU operations, with the results of the second one being added to the input so as to produce the output (Alom et al., 2018). Similarly to UNet, in the contracting path, a 2×2 max-pooling is applied in each Conv layer, whereas in the expanding path symmetrical structured blocks connected

[1] Cloud to Street, Microsoft, Radiant Earth Foundation. (2022). A global flood events and cloud cover dataset (Version 1.0). Radiant MLHub. https://doi.org/10.34911/rdnt.oz32gz.

Table 1
Trainable parameters of deep learning methods.

Training parameters	LinkNET	UNet	R2UNet	DeepLabV3+	Attention UNet	Attention R2UNet
Epochs	100					
Batch size	32					
Batch normalization	After each convolutional block					
Optimizer	ADAM					
Learning Rate (LR)	0.001					
Trainable parameters	348,374	1,965,609	2,076,665	8,572,131	2,337,933	2,448,989
Early stoping	patience = 10					
Loss function	Sparse Categorical Crossentropy					
Reduce LR On Plateau	patience = 10, min_LR = 0.0005					
Train/Val/Test split	70/10/20					

through skip connections with the contracting path are used. Note that R2UNet has 16 filters, being doubled after every recurrent residual Conv block.

(C) Attention UNet: It is a variant of the UNet that incorporates attention mechanisms to improve its segmentation capabilities through attention gates (Oktay et al., 2018). The contracting path is identical to that of the UNet architecture, but, in the expanding path the plain skip connections are replaced with attention gates, used to emphasize relevant information during both the encoding and decoding phases, mitigating the redundant computation of low-level features (Oktay et al., 2018). The Attention UNet has 16 base number of filters, doubled with each Conv block.

(D) Attention R2UNet: This architecture combines both the residual blocks and the attention mechanisms in a single architecture (Zuo et al., 2021). Specifically, the traditional Conv blocks are replaced with residual recurrent ones across the entire architecture enhancing feature extraction and avoiding the degradation learning problem during training. Considering the expanding path, attention gates replace skip connections resulting in improved learning of spatial information from the input feature maps. Similarly, with the other UNet variants, Attention R2UNet has 16 base number of filters.

(E) DeepLabV3+: Is a semantic segmentation architecture, integrating encoder–decoder processing, accurately labeling each pixel in an image with the corresponding semantic class (Chen et al., 2018). The encoder gradually reduces the feature maps capturing higher semantic information, whereas the decoder gradually recovers the spatial information. The encoding process utilizes atrous convolutions, that is a sequential pass of the input image through 1×1 Conv, two 3×3 Convs with rate 6 and 12 respectively followed by image pooling. Afterwards, the result is passed through a 1×1 Conv, up-sampled by 4 and concatenated with the low-level features from the atrous Conv, all applied during the decoding process. Finally, the result is passed through a 3×3 Conv and is upsampled by 4. Note that DeeplabV3+ uses ResNet50 as backbone during training.

(F) LinkNet: It falls in the category of encoding–decoding CNN architectures (Chaurasia and Culurciello, 2017). The contracting path involves a sequential pass of the input image through a 7×7 Conv with stride 2, a max-pooling layer of window size 2×2 and stride of 2, along with 4 encoding blocks, where each one is realized using two Conv 3×3 layers followed by max pooling. Note that batch normalization is applied between each convolutional layer, followed by ReLu to address non-linearities inside the residual unit. The expanding path reconstructs the original spatial resolution by up-sampling the features. Skip connections are incorporated to fuse low-level and high-level features, allowing for the preservation of fine-grained details and improved segmentation accuracy. Within the decoding process, each decoding layer's input is linked to its corresponding encoding one, facilitating the integration of spatial information. The decoder incorporates Conv modules with batch normalization and ReLU as well. To improve feature extraction we used ResNet-34 as backbone during training.

Machine Learning Method: For efficient segmentation of a land scenery, the K-means clustering ML algorithm is used, for a range of classes equal to $k = 5$–7. Due to the K-means unsupervised property, the land uses with common spectral signatures are grouped automatically.

3.3. Metrics

The proposed models were evaluated on both datasets with the following classification metrics: (1) Accuracy (2) Precision (3) Recall (4) F1 score,

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}, \tag{3}$$

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{4}$$

$$\text{Recall} = \frac{TP}{TP + FN}, \tag{5}$$

$$\text{F1} = 2\left(\frac{Precision \cdot Recall}{Precision + Recall}\right). \tag{6}$$

where definitions TP = True Positive, TN = True Negative, FP = False Positive and FN = False Negative, are defined from the confusion matrix. F1 score is calculated as the harmonic mean of Precision and Recall metrics. In image segmentation apart from

Table 2
Model classification accuracy using the C2S-MS Floods dataset.

Model	Accuracy	Precision	Recall	F1	IoU
B8A–B11–B12 (20 m per pixel)					
LinkNET	90.39	92.03	90.39	90.15	84.75
UNet	90.37	91.91	90.37	89.84	84.57
R2UNet	90.55	92.42	90.55	90.27	85.01
DeepLabV3+	89.87	91.77	89.87	89.60	84.09
Attention UNet	89.95	92.05	89.95	89.57	84.21
Attention R2UNet	**91.52**	**92.61**	**91.52**	**91.30**	**86.15**
B08–B04–B03 (10 m per pixel)					
LinkNET	**92.70**	**94.16**	**92.70**	**92.69**	**88.35**
UNet	89.86	91.30	89.86	89.00	83.49
R2UNet	91.63	93.26	91.63	91.43	86.48
DeepLabV3+	91.35	93.24	91.35	91.18	86.33
Attention UNet	90.50	92.72	90.50	90.17	84.95
Attention R2UNet	90.60	92.32	90.60	90.12	84.97

the 4 classification metrics, we consider the Intersection over Union (IoU) a metric that quantifies the similarity and diversity of 2 sets. It is calculated as the size of the intersection divided by the size of the union of two sets A and A′, where A is the predicted image and the ground truth label is A′. IoU is defined as:

$$\mathrm{IoU}(A, A') = \frac{|A \cap A'|}{|A \cup A'|} = \frac{|A \cap A'|}{|A| + |A'| - |A \cap A'|}. \tag{7}$$

4. *C2A-DC* framework applicability in two real-world case scenarios

In this section, we use the proposed *C2A-DC* framework to tackle two real-world environmental monitoring tasks. The first one corresponds to cloud and shadow detection tasks, whereas the second one corresponds to deforestation monitoring.

4.1. Model performance

In Table 2 the performance of the DL models for both experiments in the C2S-MS Floods dataset is cited, while in Fig. 4 the models segmentations are graphically illustrated. In the first set of experiments corresponding to the 20 m resolution per pixel, it can be observed that Attention R2UNet results in the highest accuracy value, followed by R2UNet being 1% (in subtraction terms) less accurate. Similar results are observed for the recall and F1-Score metric, whereas for the precision metric Attention R2UNet and R2UNet have close values. One can conclude here that the combination of attention and residual blocks results in better performance compared to the rest models in the 20 m resolution per pixel.

In the second set of experiments corresponding to the 10 m resolution per pixel, LinkNET results in the highest performance in all metrics, while R2UNet comes second. Moreover, it can be observed that LinkNET results in 1.18% accuracy, 1.55% precision, 1.18% recall, 1.39% F1-score and 2.20% IoU higher values in the second set of experiments compared to the first one. This verifies that the improvement in the resolution per pixel also improves the cloud and shadow detection tasks. Yet, one should not neglect the results obtained in the first set of experiments as they can be important in cases where lower resolution per pixel bands are not available.

4.2. Case study 1: Cloud and shadow detection

The proposed *C2A-DC* framework is used in a real-world case study, which is that of the 15 November 2017 flood in Mandra (Diakakis et al., 2019), one of the deadliest floods that took place in the last 40 years in Greece (Kanellopoulos et al., 2020). In this case study, the EO DC is suitable for damage assessment from the flood, which is done by selecting time stamps before the flood event, namely 5-11-2017 and 10-11-2017 and after the flood event, namely 17-11-2017 and 22-11-2017, all formed of Sentinel-2 images using $K = 13$ bands. The date format used follows DD-MM-YY. The EO DC obtained from the proposed framework contains maps with clouds and their shadows corresponding to three classes, including background = 0, cloud = 1 and shadow = 2, clustered maps for $k = 5$–7 classes and remote sensing indices including NDVI, NDWI, NDBI, SAVI and EVI. In Fig. 5 an example of the four cases before and after the flood event is shown, along with clustering using $k = 5$ classes and the NDVI.

The selected images shown in Fig. 5, are used to show three distinct phases within the same scenery of the event; (I) small clouds combined with shadows and background, corresponding to case (a) which is several days before flood, (II) images filled with clouds and shadows, corresponding to cases (b) and (c) right before and after the event respectively and (III) non-existent clouds and shadows corresponding to case (d), days after the event. Moreover, it can be observed that clouds and shadows are correctly classified into their respective classes through the phases of the event, supported by the changes in the k-means clustering values as well as the NDVI.

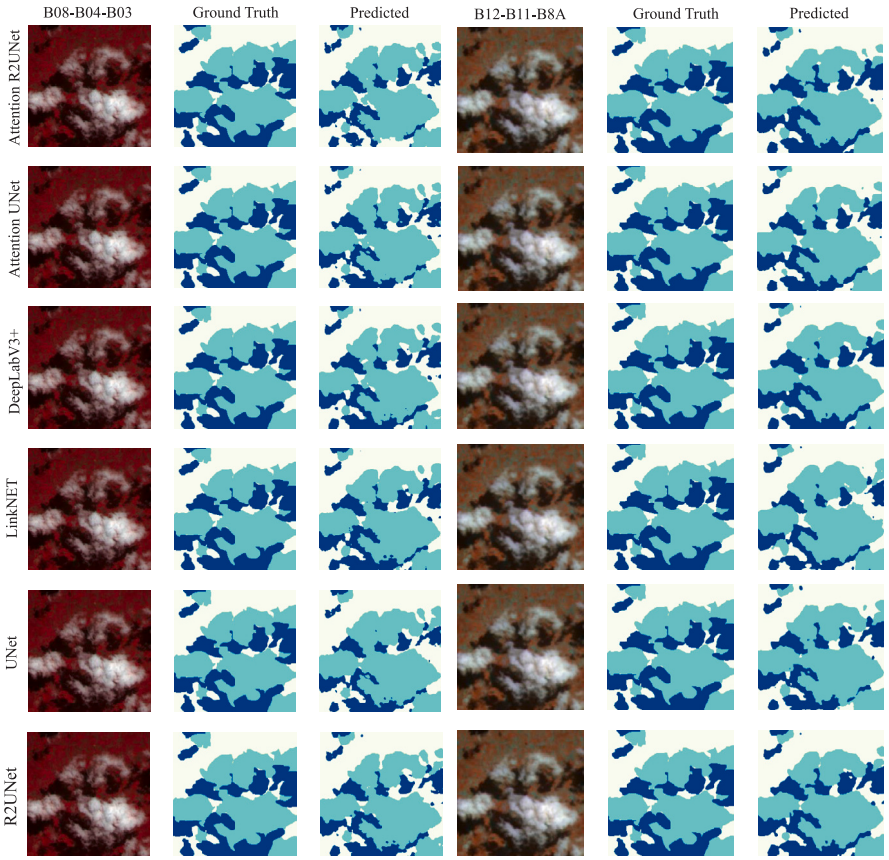

Fig. 4. Performance results on cloud and shadow detection task in two sets of experiments. First one includes B08–B04–B03 bands and corresponds to a 10-m resolution per pixel. Second one includes B8A–B11–B12 bands and corresponds to 20 m resolution per pixel.

4.3. Case study 2: Deforestation monitoring

To further showcase the effectiveness of the proposed *C2A-DC* framework, we use it for deforestation monitoring. In particular, we investigate one of the largest fires that occurred in Mati, Attica, Greece on 23-07-2018 (DD-MM-YY) (Efthimiou et al., 2020). *C2A-DC* framework forms an EO DC based on Sentinel-2 data that is used to assess the damage from the fire starting from 01-07-2023 to 30-08-2023. We set a baseline and an analysis time period to compare the spectral index in each one of those two; the baseline time period starts before the fire between 01-07-2018 and 20-07-2018, whereas the analysis starts after the fire between 01-08-2018 and 30-08-2018. For these two time periods, we select the median value of cloud-free pixels for all spectral bands in the time series forming a median mosaic. To find the significant land changes from fire, NDVI is utilized in both time periods to effectively capture vegetation loss. This is achieved by subtracting NDVI median mosaics before (baseline) and after (analysis) the event.

In Fig. 6 the post-disaster assessment is illustrated. In the first row, images illustrate the land scenery in RGB before and after the fire. In the second row, vegetation change is depicted as the difference between the two calculated NDVI median mosaics. Color Red describes vegetation loss while greener colors are for higher vegetation values. With the lower right plot, we enhance the fire footprint by visualizing values below 0, where NDVI from the analysis layer (after the fire) has also negative ones. Note that the above NDVI values are stored within the DC making the proposed *C2A-DC* framework suitable for post-disaster assessment.

5. Conclusion

In this work, the *C2A-DC* framework was proposed for efficient environmental monitoring from EO data. *C2A-DC* improves the applicability of DCs as it includes AI algorithms and data harmonization, not feasible with existing approaches. Experimental results using several CNN architectures on the C2S-MS Floods dataset showcased its effectiveness in cloud and shadow detection, an important task in flood events, while RS product calculations highlighted the importance of image time-series within the DC in deforestation monitoring. A potential extension of the framework is the inclusion of explainable AI methods so as to offer a better

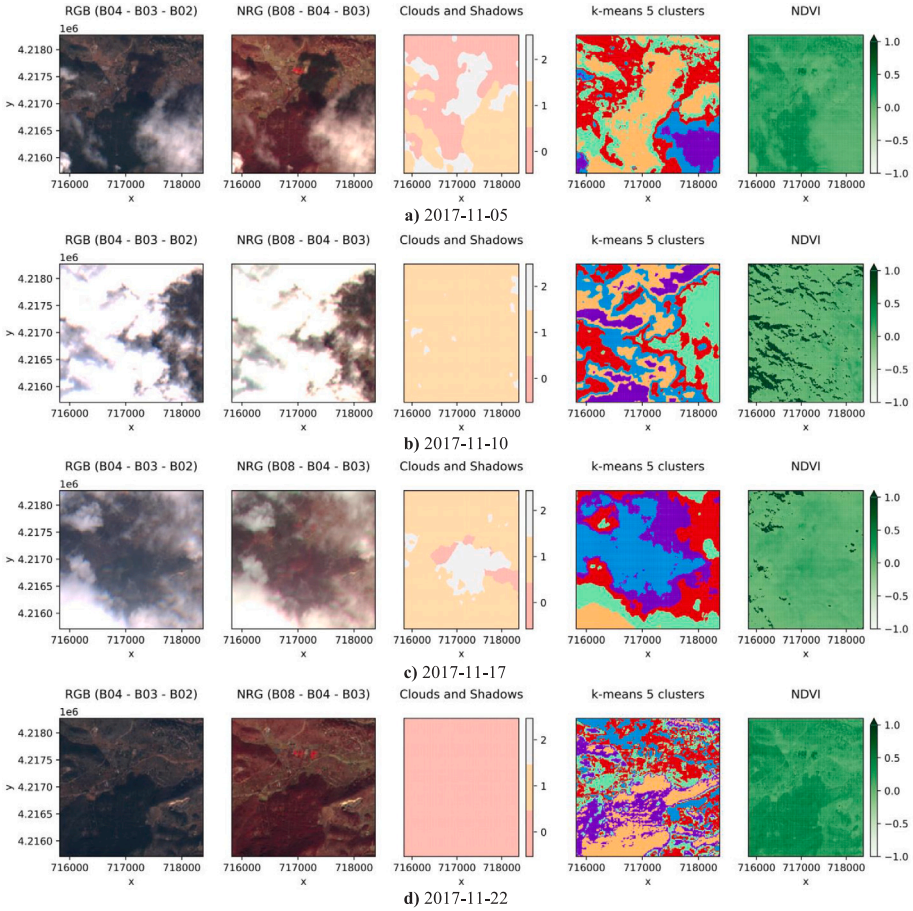

Fig. 5. Cloud and shadow detection for flood damage assessment using the proposed *C2A-DC* framework. Flood event occurred in Mandra, Attica, Greece on 15 November 2017. Date values (a) and (b) correspond to days before the event, whereas (c) and (d) to days after the event. Clouds and shadows are correctly classified by the CNN architectures. The K-means clustering algorithm and the NDVI product demonstrate the land changes.

understanding of the factors affecting the classification tasks within the scope of environmental monitoring as well as improve the trustworthiness of the AI towards the stakeholders.

CRediT authorship contribution statement

Anastasios Temenos: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Writing – original draft. **Nikos Temenos:** Conceptualization, Methodology, Supervision, Writing – original draft, Writing – review & editing. **Ioannis N. Tzortzis:** Methodology, Software. **Ioannis Rallis:** Project administration, Supervision, Writing – review & editing. **Anastasios Doulamis:** Funding acquisition, Supervision. **Nikolaos Doulamis:** Funding acquisition, Supervision.

Declaration of competing interest

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

Data availability

Data are open access from sentinel-2 satellite.

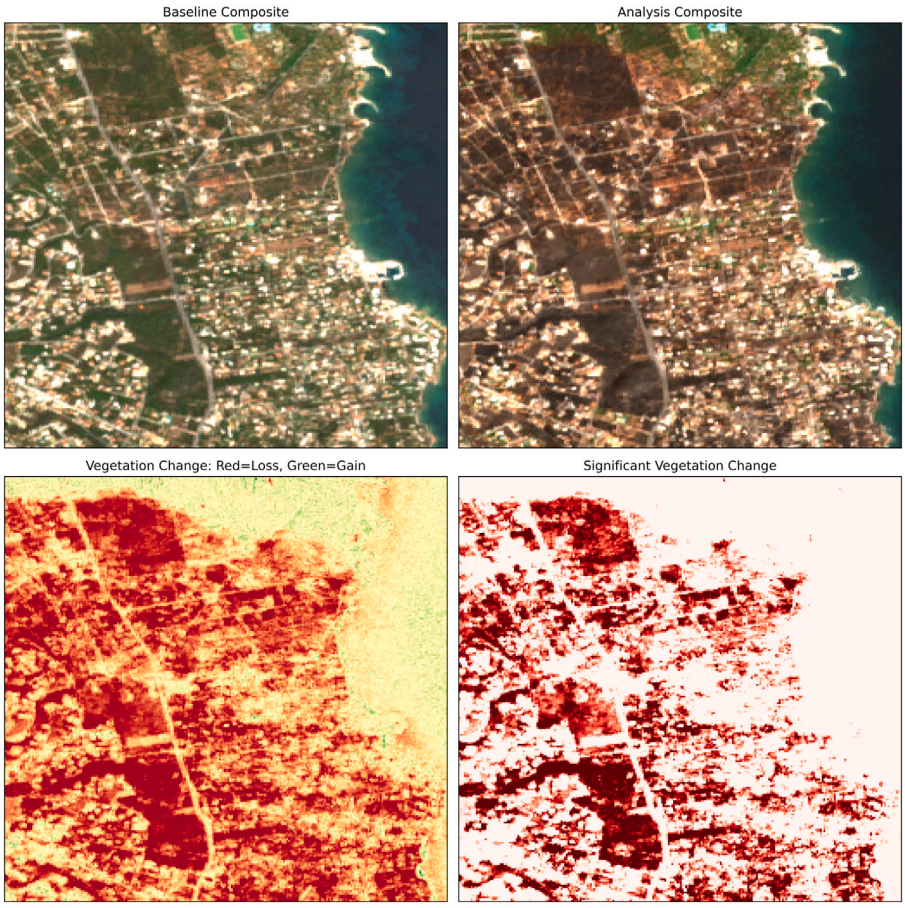

Fig. 6. Deforestation monitoring using the proposed *C2A-DC* framework. Fire event occurred in Mati, Attica, Greece on 23-07-2018. In upper left, the baseline composite corresponds to dates starting from 01-07-2018 to 20-07-2018, whereas in upper right, the analysis composite corresponds to dates starting from 01-08-2018 to 30-08-2018. Below, bottom left and right are the vegetation changes (red = loss) of their respective composites. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Acknowledgments

This work was financially supported by the Horizon 2020 European Union project HARMONIA "Development of a Support System for Improved Resilience and Sustainable Urban areas to cope with Climate Change and Extreme Events based on GEOSS and Advanced Modelling Tools" funded under the call H2020-LC-CLA-19-2020, with grant agreement No. 101003517.

References

Adger, W.N., Crépin, A.-S., Folke, C., Ospina, D., Chapin, F.S., Segerson, K., Seto, K.C., Anderies, J.M., Barrett, S., Bennett, E.M., et al., 2020. Urbanization, migration, and adaptation to climate change. One Earth 3 (4), 396–399.

Afrin, S., Chowdhury, F.J., Rahman, M.M., 2021. COVID-19 pandemic: rethinking strategies for resilient urban design, perceptions, and planning. Front. Sustain. Cities 3, 668263.

Alom, M.Z., Hasan, M., Yakopcic, C., Taha, T.M., Asari, V.K., 2018. Recurrent residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation. arXiv:1802.06955.

Bush, J., Doyon, A., 2019. Building urban resilience with nature-based solutions: How can urban planning contribute? Cities 95, 102483.

Chatenoux, B., Richard, J.-P., Small, D., Roeoesli, C., Wingate, V., Poussin, C., Rodila, D., Peduzzi, P., Steinmeier, C., Ginzler, C., et al., 2021. The Swiss data cube, analysis ready data archive using earth observations of Switzerland. Sci. Data 8 (1), 295.

Chaurasia, A., Culurciello, E., 2017. Linknet: Exploiting encoder representations for efficient semantic segmentation. In: 2017 IEEE Visual Communications and Image Processing. VCIP, IEEE, pp. 1–4.

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the European Conference on Computer Vision. ECCV, pp. 801–818.

Diakakis, M., Andreadakis, E., Nikolopoulos, E., Spyrou, N., Gogou, M., Deligiannakis, G., Katsetsiadou, N., Antoniadis, Z., Melaki, M., Georgakopoulos, A., et al., 2019. An integrated approach of ground and aerial observations in flash flood disaster investigations. The case of the 2017 Mandra flash flood in Greece. Int. J. Disaster Risk Reduct. 33, 290–309.

Efthimiou, N., Psomiadis, E., Panagos, P., 2020. Fire severity and soil erosion susceptibility mapping using multi-temporal Earth Observation data: The case of Mati fatal wildfire in Eastern Attica, Greece. Catena 187, 104320.
Ferreira, K.R., Queiroz, G.R., Vinhas, L., Marujo, R.F., Simoes, R.E., Picoli, M.C., Camara, G., Cartaxo, R., Gomes, V.C., Santos, L.A., et al., 2020. Earth observation data cubes for Brazil: Requirements, methodology and products. Remote Sens. 12 (24), 4033.
Giuliani, G., Chatenoux, B., Benvenuti, A., Lacroix, P., Santoro, M., Mazzetti, P., 2020a. 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), 3–22.
Giuliani, G., Chatenoux, B., Piller, T., Moser, F., Lacroix, P., 2020b. Data Cube on Demand (DCoD): Generating an earth observation Data Cube anywhere in the world. Int. J. Appl. Earth Obs. Geoinf. 87, 102035.
Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27.
Kanellopoulos, T.D., Karageorgis, A.P., Kikaki, A., Chourdaki, S., Hatzianestis, I., Vakalas, I., Hatiris, G.-A., 2020. The impact of flash-floods on the adjacent marine environment: The case of Mandra and Nea Peramos (November 2017), Greece. J. Coast. Conserv. 24, 1–17.
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. IEEE, pp. 5646–5649.
Kirikkaleli, D., Sowah, J.K., 2020. A wavelet coherence analysis: nexus between urbanization and environmental sustainability. Environ. Sci. Pollut. Res. 27, 30295–30305.
Leal Filho, W., Balogun, A.-L., Olayide, O.E., Azeiteiro, U.M., Ayal, D.Y., Muñoz, P.D.C., Nagy, G.J., Bynoe, P., Oguge, O., Toamukum, N.Y., 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. Sci. Total Environ. 692, 1175–1190.
Lewis, A., Oliver, S., Lymburner, L., Evans, B., Wyborn, L., Mueller, N., Raevksi, G., Hooke, J., Woodcock, R., Sixsmith, J., et al., 2017. The Australian geoscience data cube—foundations and lessons learned. Remote Sens. Environ. 202, 276–292.
Li, J., Wu, Z., Hu, Z., Jian, C., Luo, S., Mou, L., Zhu, X.X., Molinier, M., 2021. A lightweight deep learning-based cloud detection method for Sentinel-2A imagery fusing multiscale spectral and spatial features. IEEE Trans. Geosci. Remote Sens. 60, 1–19.
Mi, Z., Guan, D., Liu, Z., Liu, J., Viguié, V., Fromer, N., Wang, Y., 2019. Cities: The core of climate change mitigation. J. Clean. Prod. 207, 582–589.
Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., et al., 2018. Attention u-net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999.
Olabi, A., Abdelkareem, M.A., 2022. Renewable energy and climate change. Renew. Sustain. Energy Rev. 158, 112111.
Parker, D.E., 2010. Urban heat island effects on estimates of observed climate change. Wiley Interdiscip. Rev. Clim. Change 1 (1), 123–133.
Pour, S.H., Abd Wahab, A.K., Shahid, S., Asaduzzaman, M., Dewan, A., 2020. Low impact development techniques to mitigate the impacts of climate-change-induced urban floods: Current trends, issues and challenges. Sustainable Cities Soc. 62, 102373.
Poussin, C., Guigoz, Y., Palazzi, E., Terzago, S., Chatenoux, B., Giuliani, G., 2019. Snow cover evolution in the Gran Paradiso National Park, Italian alps, using the earth observation data cube. Data 4 (4), 138.
Ravindra, K., Rattan, P., Mor, S., Aggarwal, A.N., 2019. Generalized additive models: Building evidence of air pollution, climate change and human health. Environ. Int. 132, 104987.
Reza, M.S., Sabau, G., 2022. Impact of climate change on crop production and food security in Newfoundland and Labrador, Canada. J. Agric. Food Res. 100405.
Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. Springer, pp. 234–241.
Shen, M., Huang, W., Chen, M., Song, B., Zeng, G., Zhang, Y., 2020. (Micro) plastic crisis: un-ignorable contribution to global greenhouse gas emissions and climate change. J. Clean. Prod. 254, 120138.
Shivanna, K., 2022. Climate change and its impact on biodiversity and human welfare. Proc. Indian Nat. Sci. Acad. 88 (2), 160–171.
Tamiminia, H., Salehi, B., Mahdianpari, M., Quackenbush, L., Adeli, S., Brisco, B., 2020. Google Earth Engine for geo-big data applications: A meta-analysis and systematic review. ISPRS J. Photogramm. Remote Sens. 164, 152–170.
Temenos, A., Tzortzis, I.N., Kaselimi, M., Rallis, I., Doulamis, A., Doulamis, N., 2022. Novel insights in spatial epidemiology utilizing explainable AI (XAI) and remote sensing. Remote Sens. 14 (13), 3074.
Tsatsaris, A., Kalogeropoulos, K., Stathopoulos, N., Louka, P., Tsanakas, K., Tsesmelis, D.E., Krassanakis, V., Petropoulos, G.P., Pappas, V., Chalkias, C., 2021. Geoinformation technologies in support of environmental hazards monitoring under climate change: An extensive review. ISPRS Int. J. Geo-Inf. 10 (2), 94.
Zuo, Q., Chen, S., Wang, Z., 2021. R2AU-Net: attention recurrent residual convolutional neural network for multimodal medical image segmentation. Secur. Commun. Netw. 2021, 1–10.