AI Techniques for Near Real-Time Monitoring of Contaminants in Coastal Waters on Board Future Φsat-2 Mission

Razzano, Francesca; Di Stasio, Pietro; Mauro, Francesco; Meoni, Gabriele; Esposito, Marco; Schirinzi, Gilda; Ullo, Silvia Liberata

DOI
10.1109/JSTARS.2024.3455992

Publication date
2024

Document Version
Final published version

Published in
IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

Citation (APA)
Razzano, F., Di Stasio, P., Mauro, F., Meoni, G., Esposito, M., Schirinzi, G., & Ullo, S. L. (2024). AI Techniques for Near Real-Time Monitoring of Contaminants in Coastal Waters on Board Future Φsat-2 Mission. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *17*, 16755-16766. https://doi.org/10.1109/JSTARS.2024.3455992

AI Techniques for Near Real-Time Monitoring of Contaminants in Coastal Waters on Board Future Φsat-2 Mission

Francesca Razzano, *Student Member, IEEE*, Pietro Di Stasio, *Member, IEEE*, Francesco Mauro, *Student Member, IEEE*, Gabriele Meoni, *Member, IEEE*, Marco Esposito*, Member, IEEE*, Gilda Schirinzi, *Senior Member, IEEE*, and Silvia Liberata Ullo, *Senior Member, IEEE*

***Abstract*—Differently from conventional procedures, the proposed solution advocates for a groundbreaking paradigm in water quality monitoring through the integration of satellite Remote Sensing data, Artificial Intelligence techniques, and onboard processing. While conventional procedures present several drawbacks mainly related to late intervention capabilities, the objective of what proposed is to offer nearly real-time detection of contaminants in coastal waters addressing a significant gap in the existing literature and allowing fast alerts and intervention. In fact, the expected outcomes include substantial advancements in environmental monitoring, public health protection, and resource conservation. Namely, the specific focus of our study is on the estimation of Turbidity and pH parameters, for their implications on human and aquatic health. Nevertheless, the designed framework can be extended to include other parameters of interest in the water environment and beyond. Originating from our participation in the European Space Agency OrbitalAI Challenge, this article describes the distinctive opportunities and issues for the contaminants' monitoring on the Φsat-2 mission. The specific characteristics of this mission, with the tools made available, will be presented, with the methodology proposed by the authors for the onboard monitoring of water contaminants in near real-time. Preliminary promising results are presented, along with an introduction to ongoing and future work.**

***Index Terms*—Artificial intelligence (AI), coastal water contaminants, earth observation, machine learning, onboard processing, remote sensing (RS).**

I. Introduction

THE pressing challenges arising from population growth, escalating water demands for agriculture, energy, and industry, coupled with climate change impacts, underscore the urgent necessity to meticulously monitor and evaluate trends in water resources worldwide [1], [2], [3], [4], [5], [6], [7].

This proactive approach ensures the establishment of a solid foundation for water security, guaranteeing sustainable access to safe and usable water. Effective, integrated monitoring of the water cycle's trends and variations, encompassing both quantity and quality, requires combining satellite and in situ observations, data assimilation, and models. A review of existing observational systems underscores the imperative need for a new, integrated monitoring capability dedicated to water security. The required components for such a capability already exist and can be seamlessly integrated through collaborative efforts among national observational programs as shown by Lawford at el. [8].

Furthermore, as the global population grows, particularly in developing countries where problems, such as water scarcity and quality concerns, are expected to intensify, tensions among different sectors (e.g., agriculture versus urban users) and obstacles to balancing human needs and ecological requirements are inevitable. Tragically, over one and a half million individuals face severe health issues or perish annually due to the lack of access to safe drinking water and sanitation [9]. Given the escalating pressures on water resources, monitoring assumes critical significance across spatial and temporal scales, providing a systematic and transparent approach to addressing water-related challenges.

The United Nations and their 2030 Agenda, adopted in September 2015, with the Sustainable Development Goals, outline several key points, among which ensuring access to water and sanitation for all is crucial [10].

The issue of water pollution presents a considerable challenge to both human health and the environment across numerous countries. Therefore, the importance of monitoring contaminants in water areas cannot be understated, as it is instrumental in seeking effective and meaningful solutions to address both natural and human-induced problems. The importance of the topic is so remarkable because marine environments suffer from a lack of sufficient sampling, and the prolonged effectiveness of environmental monitoring programs is impeded by inadequate funding. Given all the above, monitoring water contaminants in coastal areas is crucial, as highlighted in numerous scientific articles addressing this issue and emphasizing the main causes and effects for each of them [11], [12]. Indicators of

Received 29 March 2024; revised 12 August 2024; accepted 3 September 2024. Date of publication 9 September 2024; date of current version 25 September 2024. *(Corresponding author: Francesca Razzano.)*

Francesca Razzano and Gilda Schirinzi are with the Parthenope, 80133 Naples, Italy (e-mail: francesca.razzano002@studenti.uniparthenope.it; gilda.schirinzi@uniparthenope.it).

Pietro Di Stasio, Francesco Mauro, and Silvia Liberata Ullo are with the Department of Engineering, University of Sannio, 82100 Benevento, Italy (e-mail: p.distasio@studenti.unisannio.it; f.mauro@studenti.unisannio.it; ullo@unisannio.it).

Gabriele Meoni is with the University of Delft, 2629 Delft, Netherlands (e-mail: G.Meoni@tudelft.nl).

Marco Esposito is with the Cosine, 2361 Warmond, Netherlands (e-mail: m.esposito@cosine.nl).

Digital Object Identifier 10.1109/JSTARS.2024.3455992

water quality, such as its physical, chemical, and biological characteristics, are typically provided by gathering samples in the field and analyzing them in a lab. However, this procedure is impractical, and providing a simultaneous water quality database on a regional scale demands significant labor and time, primarily based on in situ measurements. Furthermore, conventional ground-based point-sample techniques face challenges in capturing the spatial and temporal variations in water quality, crucial for comprehensive assessment and effective waterbody management. This limitation underscores the necessity to explore innovative monitoring methods for detecting contaminants in water. It is important to note that there are numerous existing methods that leverage Remote Sensing (RS) techniques and satellite data, which will be discussed in greater detail in the following Section II. Specifically, in our approach by integrating RS techniques with in situ measurements, we can effectively monitor water contaminants. We propose a promising approach that involves implementing a monitoring process. This process, which relies on predefined acceptable thresholds for water quality parameters, can provide near real-time alerts to emerging risks. The strength of our research lies in its ability to offer timely and proactive risk management, enhancing the overall responsiveness and effectiveness of water quality monitoring systems.. This establishes an alert system aimed at safeguarding both human health and the well-being of aquatic ecosystems. Consequently, there is an increasing demand for smart detection methods that can offer valid solutions to the problem described.

The contribution of our work represents an answer to this challenge and aims to realize an innovative method for monitoring near real-time coastal water contaminants, as described ahead in this article.

Onboard applications may be optimal for early alert systems, such as volcanic eruption detection [13], vessel detection [14], cloud detection [15], and fire smoke detection [16]. Yet there are few studies on onboard marine quality such as [17] and [18]. Compared to terrestrial counterparts, in a properly designated mission, the immediacy of detection can significantly reduce latency in notification [19], [20] [13], [21]. The artificial intelligence (AI) model we designed and proposed in this article is engineered to deliver nearly real-time detection of coastal water pollutants through the utilization of satellite data and advanced Machine Learning (ML) algorithms, to be processed on board. Our innovative solution and its monitoring capabilities can be extended to encompass open seas, safeguarding economically vital assets such as fisheries and tourism. This pioneering application of AI and satellite data significantly diminishes response times for pollution detection, because by performing the processing onboard, anomalies can be identified and transmitted to the ground in real-time. This eliminates the need to wait for the satellite data to be first sent to the ground and then processed, significantly reducing the overall response times. An example of this can be found in the study conducted by [22], which discusses the importance of onboard processing for improving response times in remote sensing applications. The use of AI onboard satellite offers Governments and decision-makers a tool to ensure the protection of public health. Our solution adds value to the existing state-of-the-art (SOTA); our onboard AI-powered solution elevates environmental monitoring by furnishing real-time insights, extending coverage, and contributing to a healthier planet. Existing solutions for water quality monitoring rely on water sampling campaigns and/or involve the use of water quality equipment, leading to substantial operational resources and costs. Simultaneously, conventional approaches are limited to detecting water anomalies after their occurrence, typically when it is too late to implement preventive measures to protect public health or businesses. On the contrary, we aim to find a near real-time solution proposed as shown in Fig. 1, demonstrating how the process involves extracting parameters from an image, identifying anomalies, and transmitting only anomalous data toward the ground stations. The main contributions of this work are as follows.

1) *Onboard AI Processing:* This work pioneers the use of onboard AI processing for near real-time monitoring of water quality, specifically focusing on turbidity and pH levels. By performing data processing directly on the satellite, we significantly reduce the latency compared to traditional methods that rely on ground-based postprocessing.
2) *Integration with ΦSat-2 Mission:* The proposed methodology is implemented and tested on the hardware of the ΦSat-2 mission, part of the ESA OrbitalAI Challenge. This demonstrates the feasibility and effectiveness of our onboard AI approach in a real-world satellite mission.
3) *Improved Response Time:* Our approach enables much faster detection and response to water quality issues, which is critical for early warning systems and timely intervention in coastal areas.
4) *Scalability and Extensibility:* While this study focuses on turbidity and pH parameters, the onboard AI framework is designed to be scalable and extensible to other water quality parameters and environmental monitoring tasks.
5) *Promising Preliminary Results:* The initial results from our onboard processing framework show great promise, and we outline potential future work to further refine the model and extend its applications to other environmental monitoring challenges.
6) *Advancement in Environmental Monitoring:* By leveraging advanced AI algorithms and the processing capabilities of modern satellite missions, this work aims to advance the SOTA in environmental monitoring, contributing to public health protection and resource conservation.

The rest of this article is organized as follows. Section II illustrates a wide analysis of the SOTA for the detection of water contaminants, focusing on AI-based techniques. In Section II-A, the main parameters monitored and presented in SOTA are introduced and discussed. The choice of turbidity and pH as the first parameters to study is also justified by their relationship with many other contaminants. A description of the future Φsat-2 mission is given in Section II-B and II-C to furnish important information on its characteristics to interested researchers, including the details of the processor and the simulator that will work on board. In Section III, details on the Myriad 2 device are given. Section IV describes the dataset creation, based on in situ data of Ligurian ARPA used as Ground Truth (GT) and the Φsat-2 satellite data made available through its simulator.

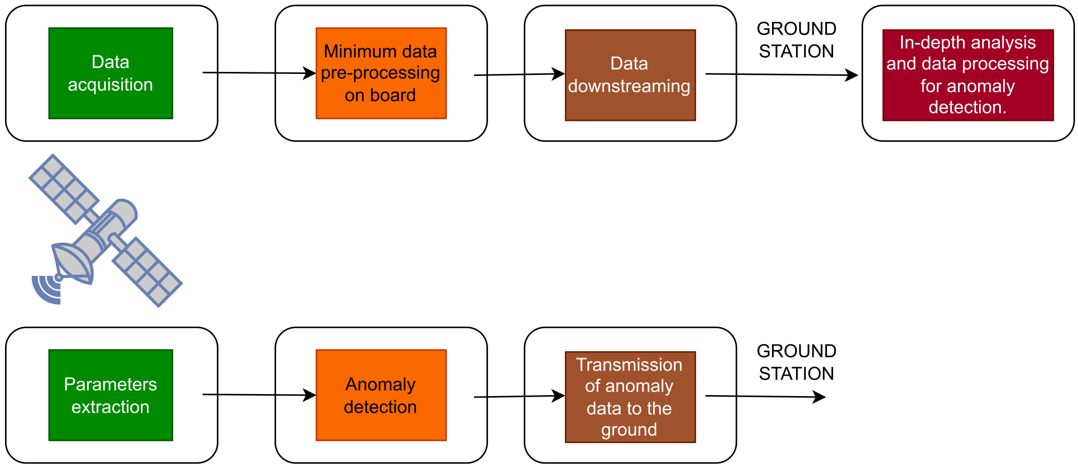


Fig. 1. Top Workflow (conventional approach): Data is acquired and minimally preprocessed on the satellite before being transmitted to the ground station, where in-depth analysis for anomalies is conducted. Bottom Workflow (our proposed procedure): The satellite extracts parameters and detects anomalies onboard, transmitting only the relevant anomaly data to the ground station, reducing data transmission and enabling faster detection.

The proposed methodology is explained in Section V, highlighting the strengths of onboard AI processing in delivering near real-time solutions for monitoring water contaminants, by filling a gap in the present literature. The models are also introduced in this section with the proposed AI for Early Detection of Environmental Threats (AI4EDoET) architecture and there is also a brief introduction to the metrics. Section VI discusses the results reported for both cases: Results with CPU in Section VI-A and results with Myriad 2 Section VI-B. Finally, Section VII concludes this article.

II. Background

The integration of RS techniques enhances water quality assessment by overcoming traditional limitations, as highlighted in [23]. RS, when combined with in situ measurements, offers significant advantages for monitoring Ocean Water Quality (OWQ) parameters on both large and coastal scales. This approach is based on the optical measurability of most water quality parameters from satellite data. The review covers OWQ parameters such as chlorophyll-a (Chl-a), colored dissolved organic matter, Secchi disk depth, turbidity, total suspended sediments (TSS), SST, and chemical oxygen demand (COD), along with monitoring models utilizing various RS datasets. Semianalytical models coupled with multivariate statistical analysis have gained attention due to their advantages over analytical approaches, which are deemed complex. Empirical and semiempirical models, though easy to implement, require sufficient in situ measurements for coefficient estimation. Efforts have been made to integrate RS datasets with alternative environmental models to address limitations. Adjovu et al. [24] emphasizes the integration of multispectral sensors for water quality monitoring, particularly beneficial in developing regions. However, challenges arise with coarse-resolution images, necessitating refinement through atmospheric correction models and exploration of high-resolution and hyperspectral imagery. In addition, the evolution of ML models and their integration into RS have introduced a novel dimension to water quality metrics.

For example, Kwong et al. [25] utilized Sentinel-2 satellite imagery and Google Earth Engine to monitor marine water quality in Hong Kong. They applied artificial neural network (ANN) models, achieving an root mean square error (RMSE) of 1.954 Nephelometric Turbidity Unit (NTU) for turbidity, RMSE of 4.513 mg/L for suspended solids (SS), and RMSE of 2.18 μg/L for Chl-a. These results highlight the effectiveness of ANN in providing continuous and updated insights into water quality. In contrast, the Cubist Model (ML-CB) used in the study proposed by Do et al. [26] combines both optical and SAR data (Sentinel-2A and Sentinel-1A) to estimate surface water pollutants such as TSS, COD, and biological oxygen demand (BOD). The accuracy of this model was validated against field survey data using regression models, demonstrating its robust performance in pollutant estimation and its potential for effective water quality management. Another significant contribution is the study proposed by Hibjur et al. [27] that applies RS-derived indices and ML algorithms to analyze spatio-temporal changes in surface water bodies in the lower Thoubal river watershed. By integrating Landsat data from multiple years and applying ML techniques, this study highlights the random forest model's superior performance in accurately extracting surface water across different seasons. Furthermore, Nazeer et al. [28] showed that ML methods, particularly neural networks (NN), surpass empirical predictive models in estimating SS and Chl-a in complex water bodies. Moreover, Vakil et al. [29] explored the use of ANN and linear regression models to determine the relationship between Landsat 8 Operational Land Imager data and total nitrogen (TN) and total phosphorus (TP) concentrations. Compared to LR models, the ANN model showed higher accuracy in the testing phase, with R values of 0.81 for TP and 0.93 for TN. In addition, Sharaf El Din et al. [30] employed a Back-Propagation Neural Network (BPNN) framework to measure surface water quality parameters, focusing mainly on local oceanographic processes. This study reported high R^2 values—ranging from 0.930 to 0.991—for turbidity, TSS, COD, BOD, and dissolved oxygen, illustrating the BPNN model's effectiveness in assessing surface water quality. Our study introduces AI onboard

TABLE I
COMMON TURBIDITY VALUES IN NATURAL ENVIRONMENTS, REPORTED IN NTU (NEPHELOMETRIC TURBIDITY UNITS) [31]

Situation	Turbididty value
Standards for drinking water by EPA	$<$ 0.3 NTU in 95% of samples; never higher than 1 NTU
Treated water	0 to 1 NTU
Fresh water with visibility exceeding 21.5 inches	$<$ 10 NTU
Fresh water with reduced visibility to 2.5 inches	240 NTU
Brief-term pressure on aquatic organisms	$>$ 10 NTU
Suboptimal conditions for the majority of aquatic organisms	$>$ 100 NTU

for contaminant detection, a pioneering approach unmatched in existing literature. This methodology not only accelerates anomaly detection, but also offers a forward-looking solution, qualifying for potential deployment in space missions like Φsat-2 or the International Space Station, following participation in the OrbitalAI Challenge endorsed by ESA Φ-lab.

TABLE II
TYPICAL PH VALUES FOUND IN NATURAL ENVIRONMENTS [36]

Situation	pH value
Acidic pH	<7
Pure water (neutral pH)	7
Alkaline (basic pH)	>7

A. Water Contaminants Parameters

With a focus on key parameters such as Temperature, Dissolved Oxygen, pH, Turbidity, Macroinvertebrates, Escherichia coli (E. coli), Nutrients, Habitat Assessment, and Metals, the rationale behind this monitoring approach is rooted in extensive research, notably drawing inspiration from the U.S. Environmental Protection Agency (EPA) [32] where the detailed reasons are described. In a particular way, turbidity, a pivotal parameter, is intricately connected to TSS, functioning as a key indicator of shared water quality characteristics. The impact of elevated turbidity levels on aquatic health is profound, affecting fish gills, visibility for predators, light penetration for aquatic plants, and fish resistance to disease. A variety of factors, both natural and human-induced, contribute to turbidity changes, underscoring the imperative need for continuous monitoring [33], [34]. Understanding typical turbidity values for various scenarios [31] is paramount as shown in Table I. Constant vigilance in monitoring turbidity is crucial for safeguarding water ecosystems and ensuring the well-being of both aquatic life and human populations. For this reason, considering thresholds established by international governmental entities facilitates the alert system process, thereby contributing to mitigating risks to human health [35]. Ensuring water quality involves vigilant monitoring of various parameters affecting turbidity, with pH standing out as a critical factor. In fact, the logarithmic nature of the pH scale emphasizes its importance, where even a slight one-unit change signifies a ten-fold shift in acidity. The pH not only influences water chemistry and toxicity, but also affects the solubility and toxicity of metals. Fluctuations in pH levels are a daily occurrence in lakes and rivers, influenced by factors like photosynthesis, respiration of aquatic plants, and human activities and in a specific way we can consider common value of pH as shown in Table II.

When https://www.fondriest.com/environmental-measurements/parameters/water-quality/ph/pH levels dip below 7.6, coral reefs are prone to collapse due to insufficient calcium carbonate. Freshwater species with heightened sensitivity, like salmon, thrive in pH levels ranging from 7.0 to 8.0. Exposure to levels below 6.0 can result for salmons in severe distress and physiological damage caused by the absorption of metals. The significance of pH monitoring extends to its impact on marine and freshwater ecosystems. In coastal areas, continuous monitoring of turbidity and pH is imperative for safeguarding both human health and the delicate balance of aquatic ecosystems. Furthermore, the inclusion of additional parameters, such as E. coli, adds another layer of importance to water quality assessment. E. coli serves as a vital indicator of fecal contamination, providing insights into the potential presence of disease-causing bacteria and viruses in freshwater. Elevated levels of E. coli pose risks to individuals engaging in recreational activities, leading to symptoms like vomiting and diarrhea [37], [38]. The interconnectedness of E. coli concentrations with various parameters like turbidity, TSS, phosphorus, nitrate, and BOD underscores the comprehensive approach needed for effective WQ management. Monitoring E. coli levels alongside these parameters is crucial for public health and environmental considerations, ensuring the safety of water bodies for recreational activities and maintaining the overall well-being of ecosystems. Having established the crucial role that monitoring contaminants in coastal waters plays in our society, the next Section V will delve into the methodology and application context of our study, which has yielded promising initial results.

B. Φsat-2 Mission

AI onboard for EO has recently gained a huge interest for the positive impacts on many monitoring applications over our planet. ESA was a pioneer in taking the initial steps in this extremely difficult area of research with the Φsat-1 satellite launched on September 3, 2020. This was the first of two 6 Units CubeSats that make up the Federated Satellite System (FSSCat) mission, also known as the FSSCat. Regarding the second mission, https://platform.ai4eo.eu/orbitalai-phisat-2 the Φsat-2 satellite, the scheduled launch date is August 2024, with an operation time planned for 10.5 months, potentially extendable by 3–8 months. This mission will integrate onboard processing capabilities, including AI, along with a Visible to Near Infra-Red (VIS/NIR) multispectral instrument capable of

TABLE III
MultiScape100 Spectral Bands

Band	Centre wavelength (nm)	FWHM Bandwidth (nm)	HPP Cut-On (nm)	HPP Cut-Off(nm)	Approximate sensor line number
# 0: PAN	625	250	500.0	750.0	1536
# 1: MS 1	490	65	457.5	522.5	2176
# 2: MS 2	560	35	542.5	577.5	1964
# 3: MS 3	665	30	650.0	680.0	1748
# 4: MS 4	705	15	697.5	712.5	1108
# 5: MS 5	740	15	732.5	747.5	896
# 6: MS 6	783	20	773.0	793.0	680
# 7: MS 7	842	115	784.5	899.5	1324

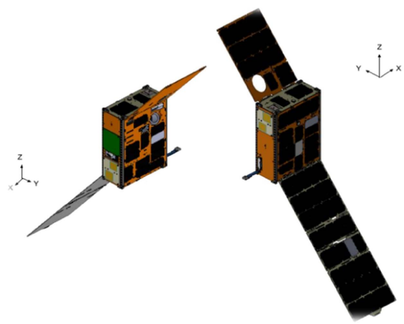


Fig. 2. This image depicts the Φsat-2 spacecraft in its deployed configuration. The spacecraft is a 6U CubeSat, constructed using the standard Open Cosmos OpenSat 6U platform, optimized for space missions requiring compact and efficient satellite solutions [39].

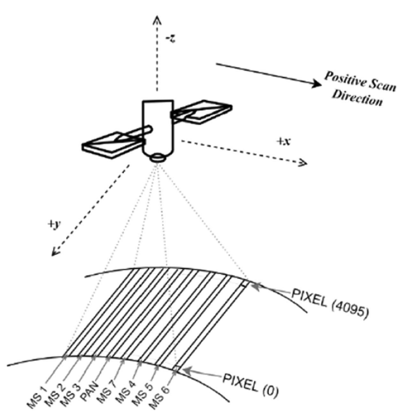


Fig. 3. Ground projection of the Φsat-2 spacecraft's imaging system. The MultiScape100 instrument, developed by Simera Innovate GmbH, is a push-broom imager that captures continuous line-scan images across 8 spectral bands in the visible and near-infrared (VNIR) range [39].

acquiring 8 bands (7 + Panchromatic) at a medium to high resolution. These functionalities will be allocated to a set of specialized applications designed to operate on the spacecraft. The satellite is a 6U CubeSat utilizing the established OpenSat 6U platform from Open Cosmos. Fig. 2 illustrates the Φsat-2 spacecraft in its deployed configuration.

Specifically, concerning the payload aspect, the MultiScape100 instrument developed by Simera Innovate GmbH operates as a push-broom imager. This device ensures continuous line-scan imaging across the 8 spectral bands within the VIS/NIR spectral range. The push-broom instrument captures images by scanning along the ground track while the spacecraft orbits the Earth, as illustrated in Fig. 3 below. This scanning process is conducted separately for each spectral band.

Detailed information on each spectral band, including its corresponding line number on the detector plane, is presented in Table III.

C. Φsat-2 Simulator

Within the framework of the OrbitalAI Challenge, a pivotal simulator for Φsat-2 has been released on GitHub [40], to assist participants in crafting authentic applications. The primary objective is to provide users with an intuitive tool proficient in realistically simulating diverse products generated on board, unfettered by geographical or temporal coverage constraints. The simulator caters to the necessity of simulating any Area of Interest (AOI) without the encumbrance of cost or commercial license restrictions, mandating the utilization of Sentinel-2 data as the primary input. The spectral and spatial attributes of Sentinel-2 furnish a robust underpinning for Φsat-2, notwithstanding disparities in spatial resolution. The spectral bands of Φsat-2 exhibit a significant higher spatial resolution (in meters) than Sentinel-2 as shown in Table IV. In addition, the available products on board Φsat-2 include those of levels L1A, L1B, and L1C. In our investigation, we employed L1C products, characterized by Top of Atmosphere Reflectance in sensor geometry, precise georeferencing, and fine band-to-band alignment (RMSE < 10 m). It is important to note that, in contrast to https://sentinel.esa.int/documents/247904/685211/Sentinel-2-Products-Specification-Document Sentinel-2 nomenclature, https://ai4eo.eu/wp-content/uploads/2023/02/Phisat-2_Mission_Overview_Web.pdf Φsat-2 products at this level are not orthorectified.

The operation of the simulator can be summarized as follows: First, the Sentinel-2 L1C bands ("B02," "B03," "B04," "B08," "B05," "B06," "B07") and the Scene Classification mask are retrieved, with additional details accessible regarding the bands. Subsequently, separate arrays are generated for the cloud, cloud shadow, and cirrus masks. Optionally, time-frames may be filtered based on data coverage, and metadata concerning solar irradiance and Earth–Sun distance are then acquired. Radiances are computed from reflectances, and a pan-chromatic image is generated through a linear combination of the input bands. Spatial resampling is conducted to align with the pixel size of Φsat-2 and band-to-band misalignment at the L1A/L1B level is simulated, along with signal degradation attributable to Signal-to-Noise Ratio and Module Transfer Function. For L1C, reflectances are simulated from radiance values, and the AOI is partitioned into a more compact set of image chips. Finally,

TABLE IV
COMPARISON BETWEEN THE MULTISPECTRAL PAYLOADS OF SENTINEL-2 AND Φ SAT-2 AT A HIGH LEVEL, ASSUMING AN ORBITAL HEIGHT OF 500 KM AND A CONSISTENT GROUND SAMPLING DISTANCE (GSD) OF 4.75 M ACROSS ALL EIGHT BANDS

Sentinel-2				Φsat-2			
Id	**Spatial resolution [m]**	**Central wavelength [nm]**	**Bandwidth [nm]**	**Id**	**Spatial resolution [m]**	**Central wavelength [nm]**	**Bandwidth [nm]**
2	10	492.4	66	MS 1	4.75	490	65
3	10	559.8	36	MS 2	4.75	560	35
4	10	664.6	31	MS 3	4.75	665	30
5	20	704.1	15	MS 4	4.75	705	15
6	20	740.5	15	MS 5	4.75	740	15
7	20	782.8	20	MS 6	4.75	783	20
8	10	832.8	106	MS 7	4.75	842	115
				PAN	4.75	625	250

bands and masks for these image chips are saved, culminating in the creation of an AI-ready dataset.

III. MYRIAD 2 DEVICE

The AI processing engine integrated into Φsat-2 is based on the Myriad 2 Vision Processing Unit (VPU). The https://ark.intel.com/content/www/us/en/ark/products/122461/intel-movidius-myriad-2-vision-processing-unit-4gb.htmlMyriad 2 VPU is a specialized hardware accelerator designed by Intel/Movidius for handling computer vision and deep neural network tasks. It is part of the Myriad family of VPUs and is specifically tailored for applications such as image and video processing, object recognition, and other vision-related tasks. This multifaceted device, with its compact dimensions, exemplifies a sophisticated and versatile solution for a range of AI applications [41].

Several studies investigate the capabilities of Myriad 2, comparing it with other solutions and applying it in various contexts. It has flight heritage with Φsat-1 and is energy-efficient, with a low development time [42]. The benefits of using Myriad 2 over traditional field programmable gate arrays and CPUs are assessed by the authors in [43]. Other research works, like [44], holds significant importance in this context introducing convolutional neural network (CNN) accelerators using Myriad 2 through two design approaches: 1) deploying CNNs on a power-efficient System on Chip (SoC); and 2) implementing a VHDL application-specific design with a corresponding FPGA architecture. Both systems aim to optimize time performance for specific dataset applications. In current advancements, additional benefits are investigated within the state-of-the-art framework that focuses on enabling an efficient Support Vector Machines implementation on an ultra-low-power multicore SoC, specifically the Intel/Movidius Myriad 2 [45].

IV. DATASET CREATION

To create an onboard AI system for the detection of water pollutants using satellite RS data, we built a regression model considering the Φsat-2 spectral bands as independent variables and the in situ data measurements as dependent variables. In this section we present a suitable dataset created to perform our task. For in situ data measurements (chosen as dependent variables), an open-access database is made available by https://www.arpal.liguria.it/ARPA Liguria containing measurements of the quality of the Ligurian Sea in waters, sediments, and marine organisms for the verification of the health status of the coastal

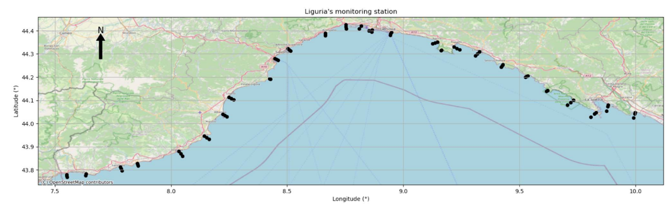


Fig. 4. Distribution of the 76 monitoring stations in the Liguria region, collecting multiple data points daily, with contaminant levels assessed at the sea surface to construct the ground truth (GT).

ecosystem and the monitoring of bathing waters. A monitoring network consisting of stable detection points enables retrieval of different parameters, belonging to different environmental matrices, that are periodically analyzed: Water, plankton, sediments, and benthonic biocenoses. The sampling points, called stations, are identifiable through a description that shows in succession the name of the municipality of belonging, the location name, the distance from the coast, and the environmental matrix under investigation. Since 2007, the Liguria coastal waters have been divided into 26 macroarea where the waters, sediments, and animal and plant populations are periodically analyzed (see Fig. 4)[1].

We consider 76 different stations which are located 500, 1000, and 3000 m from the coast. For each monitoring station, different dates of acquisition for the parameter under test are present, and these dates appear repeated several times because, on the same day, different acquisitions were made at different heights of the sensor. In our case, the value of the contaminants has been assessed at the sea surface level to construct the GT. The ground data points are designed to match the satellite data for accurate comparison and analysis. Consequently, their number corresponds to the number of satellite data. For the creation of the dataset, the first step involved selecting in situ data of to the Liguria region. Subsequently, the identified data were downloaded from the Arpa database. After that, we moved to the simulator Python API and we adjusted the bounding box size from 20140 to 1216 m, creating a square around the Point of Interest (PoI). We considered only a restricted area surrounding

[1]The visualization of Fig. 4 integrates a base map from OpenStreetMap (https://www.openstreetmap.org/#map=8/43.894/4.890), overlaid using the Python library *contextily* with the *OpenStreetMap Mapnik* provider. The graph was constructed by the authors with *pandas* for data manipulation, *GeoPandas* for geospatial operations, and *matplotlib* for graphical rendering. *Shapely* was used for handling geometric objects, facilitating accurate spatial representation of the data.

the PoI, generating a single patch of dimensions 256 × 256 pixels (considering a spatial resolution of 4.75 m). Reducing a patch to 256 × 256 pixels is also done to accommodate the limited memory requirements of accelerators. It is typical for edge applications [46]. Specifically, Φsat-2 data temporally and spatially aligned with the in situ data were selected (a temporal tolerance of 3 days has been taken into account, considering the revisit time of Φsat-2). The implemented code, which will be made available after publication, allows for automated downloading of additional Φsat-2 data temporally and spatially aligned with new in situ data. The result is a dataset of 1805 unique samples defined by latitude and longitude coordinates, where each of these areas covers a spatial expanse of 1216 km × 1216 km.

$$X_{lat,lon} \in R^{W \cdot H \cdot B} \tag{1}$$

where *lat* and *lon* are the center coordinates, *W* the width of the image (256 px), *H* the height of the image (256 px), and *B* the number of bands of the image (8). The data processing level chosen is L1C (without atmospheric correction). We used L1C data instead of atmospherically corrected data because the correction process is computationally intensive to perform on board. Furthermore, as shown in Razzano et al. [47], using atmospherically corrected data does not yield significant improvements.

V. Method Proposed

Since punctual data are available, we started from a Regression model, which is fully connected and subsequently, we parallelized it, by transferring the information to a CNN that allowed us to calculate the regression efficiently. The final network is a CNN that has not been directly trained, but whose weights have been transferred by a regression neural network. The proposed network has been defined as the AI4EDoET. The proposed method has been implemented in Python through PyTorch and the code will be made available after publication.

1) Fully Connected Neural Network Regression Model: Our starting AI model is a *Fully Connected Neural Network for regression*, an ANN specifically designed to solve regression problems. The goal is to predict pH and turbidity values, given Φsat-2 bands as input.

2) AI4EDoET: Our AI4EDoET network is CNN-based model and processes input image with a shape of 256 × 256 × 7. It produces as output a 25 × 25 matrix. Each output matrix element contains estimates of one water pollutant over 10 × 10 pixels of the input patch. The first convolutional layer performs spatial averaging over a 10 × 10 window. In the regressor network, averaging over a 10 × 10 pixel window was performed as a preprocessing step. In our CNN, this operation is handled by the first layer, which emulates the averaging. It is worth noting that the choice of window size is justified by the fact that initially, on Sentinel-2 data with a spatial resolution of 10 m, averaging over 4 pixels was performed. Now, with PhiSat-2 data, which has a resolution of 4.75 m, we have transitioned to averaging over 10 pixels. Moreover with this choice of the window size all satellite data corresponding to the 76 ground stations are situated in the sea. The nth convolutional layer $(n > 1)$ maps the $(n-1)$th fully-connected layer in the corresponding regression network by utilizing 1 × 1 kernels with no spatial averaging. In Fig. 5 it has shown that the network has N+1 layers, where N is the starting regression network layer number, and +1 indicates that the first layer is a convolutional layer that emulates the spatial average made to calculate the data to train this network. The number of output channels is set equal to the number of neurons in the corresponding fully connected layer, while the number of input channels matches the number of neurons in the preceding fully connected layer.

We adopted two different models: one for the estimation of pH and one for the Turbidity (the hyparameters setting is the same, but the output weights are different). The entire workflow of our proposed approach is schematized in Fig. 6. Since on board, the entire Φsat-2 tile is available, the initial operation involves tiling the image, resulting in single patches that can be processed by the CNN. Subsequently, binary maps are generated by applying a specified threshold, classifying values as either 0 or 1, based on the alert value of the contaminant. The final step involves reconstruction through a de-mosaicking process, aimed at restoring the complete input tile of Φsat-2. However, the demosaicking part was not considered in our work because, in the subsequent phase of the mission launch, it will be implemented by other companies. In the context of our participation in OrbitalAI Challenge, Ubotica has provided us the access to Rupia, an application designed for remote execution of models on the Myriad 2 device. Rupia introduces a virtual environment tailored for running inference models on Ubotica's Space-ready platforms. This platform allows for remote execution of Edge AI inference without the need for additional hardware. It is supported by a robust and extensively tested backend infrastructure and seamlessly integrates with existing workflows for Φsat-2. The implementation on Myriad 2 involved conversion first to ONNX and then through Rupia, with rigorous testing throughout the process. The choice of the CNN framework for our study is primarily influenced by the constraints associated with onboard implementation. Specifically, to deploy a model on the Φsat-2 satellite, it must be converted from Python to ONNX and subsequently to Rupia. The conversion process is well-supported for CNN models, which facilitated their integration into the satellite's system. Other frameworks, such as Catboost Regressor, were also tested in our research, as mentioned in Razzano et al. [47]. However, at present, integrating Catboost Regressor into the satellite system is not feasible. Therefore, to ensure compatibility and successful deployment, we opted for the CNN framework, which aligns with the current technical requirements and available support tools.

VI. Results

This section is divided into two subsections. The first subsection includes the results obtained for the estimation of Turbidity and pH values using CPU hardware. The second subsection contains the results obtained through Myriad 2 hardware. Table V provides a summary of the layers and configuration details of the Fully connected Neural Network while Table VI provides a summary of the configuration details of the CNN.

The models performances are assessed using appropriate regression metrics, such as RMSE and the Mean Absolute Error

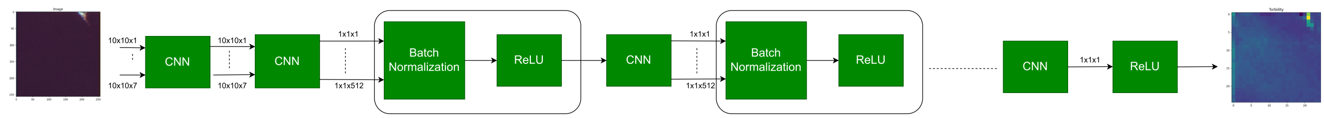


Fig. 5. AI for early detection of environmental threats (AI4EDoET)—the proposed neural network.

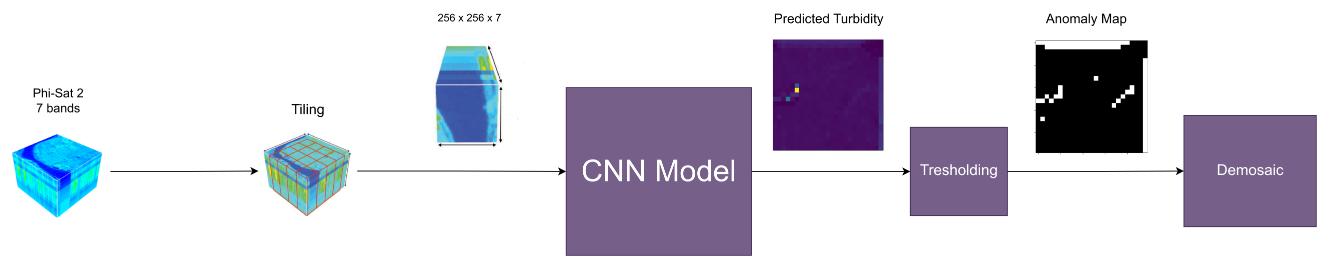


Fig. 6. Final schematic about the workflow for the implementation of AI4EDoET model on board for Turbidity estimation.

TABLE V
FULLY CONNECTED NEURAL NETWORK'S SETTINGS

Parameter	Value
Number of hidden layers	5
Number of nodes in hidden layers	[512, 512, 512, 512, 43]
Activation function	Rectified Linear Unit (ReLU)
Batch normalization	Yes
Dropout (p)	0.25
Optimizer	Adam
Loss function	Root Mean Square Error (RMSE)
Number of epochs	2000

TABLE VI
CONVOLUTIONAL NEURAL NETWORK'S SETTINGS

Parameter	Value
window size	10
input channels	7
hidden layers	[512, 512, 512, 512, 43]
output channels	1

(MAE). The choice is also based on the need for consistency and comparability with existing literature or benchmark studies, ensuring a standardized approach to evaluating and communicating the performance of models or methods [25], [48] [49]. In particular, Kwong et al. [25] leveraged Sentinel-2 satellite imagery and Google Earth Engine to automatically map and monitor marine water quality in Hong Kong, providing continuous and updated insights using ANN models. These models are optimized through cross-validation and involve a wide range of input variables and neurons selected for different parameters. Similarly, Hafeez et al. [48] compared ML algorithms to retrieve water quality indicators from satellite data in Hong Kong's complex Case-II waters, aiming to improve the accuracy and reliability of RS techniques. Their study includes ANN, Random Forest, Cubist regression (CB), and support vector regression models. Reflectance data from Landsat (5, 7, and 8) were compared with in situ reflectance measurements to assess the performance of the ML models. In addition, Pereira et al. [49] utilized cloud-based processing of Landsat images to estimate pH levels of lakes in the Brazilian Pantanal, introducing a new classification system that enhances ecological understanding and supports conservation efforts. They predict pH values by applying linear multiple regression and symbolic regression based on genetic programming (GP), a powerful ML modeling technique introduced by Koza et al. [50].

A. Results With CPU

With regard to the outcomes presented in this section, a qualitative and quantitative analysis is provided concerning Turbidity and pH parameters. The proposed model is composed of five hidden layers with [512, 512, 512, 512, 43] nodes. Several simulations were performed to set the number of nodes and hidden layers, in such a way as to minimize the root mean square error (RMSE) between the prediction and the GT.

Furthermore, the data are systematically prepared, incorporating a suitable splitting between training, testing, and validation sets with percentages of 55%, 20%, and 25%, respectively. We compare the metrics of our results with other approaches reported in the SOTA that monitor the same contaminants using alternative AI-based frameworks. However, our model not only introduces a valuable innovation in a social context by addressing the issue of onboard water monitoring via satellite in real-time, but it also yields superior results when compared to the mentioned studies against which comparisons are made. It should also be noted that for the proposed model no atmospheric correction of the data is foreseen and this represents a further advantage compared to the other methods present in the SOTA, as the atmospheric correction operation is heavy to do on board [51].

In Figs. 7 and 8 are illustrated qualitative results by depicting the distributions of predictions on the training, validation, and test sets, juxtaposed with the expected distributions of Turbidity and pH parameter (Ground Truth). As can be seen from the graphs, the estimated distributions well map the trends of the GT distributions, indicating accurate mapping by the model.

In a quantitative way, the validation metrics (RMSE and MAE) for Turbidity and pH values are reported in Tables VII and VIII. These results are compared with other approaches

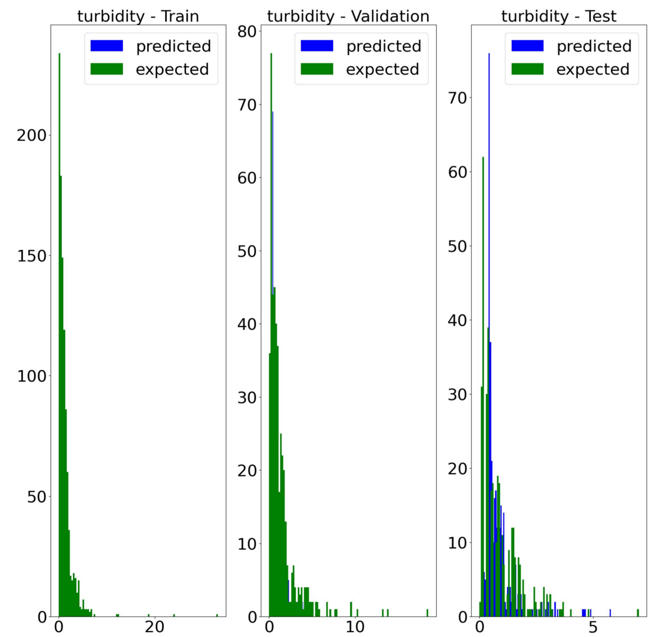


Fig. 7. Qualitative results for Turbidity distributions of predictions across the training, validation, and test sets, compared to the expected Ground Truth (GT) distributions. The graphs demonstrate that the predicted distributions align closely with the trends observed in the GT distributions.

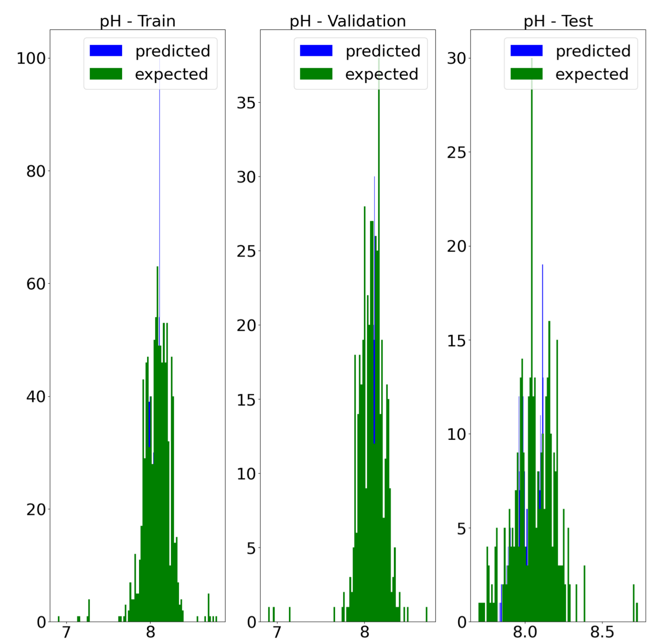


Fig. 8. Qualitative results for pH distributions of predictions across the training, validation, and test sets, compared to the expected Ground Truth (GT) distributions. The graphs demonstrate that the predicted distributions align closely with the trends observed in the GT distributions.

TABLE VII
COMPARISON OF TURBIDITY METRICS WITH OTHER STUDIES REVEALS THE SUPERIOR PERFORMANCE OF OUR ANALYSIS

Turbidity [NTU]	Methodology	RMSE	MAE
Our study	**AI4EDoET**	**1.0973**	**0.7426**
Kwong et al. [25]	ANN	1.954	1.607
Hafeez et al. [48]	ANN	3.10	2.61

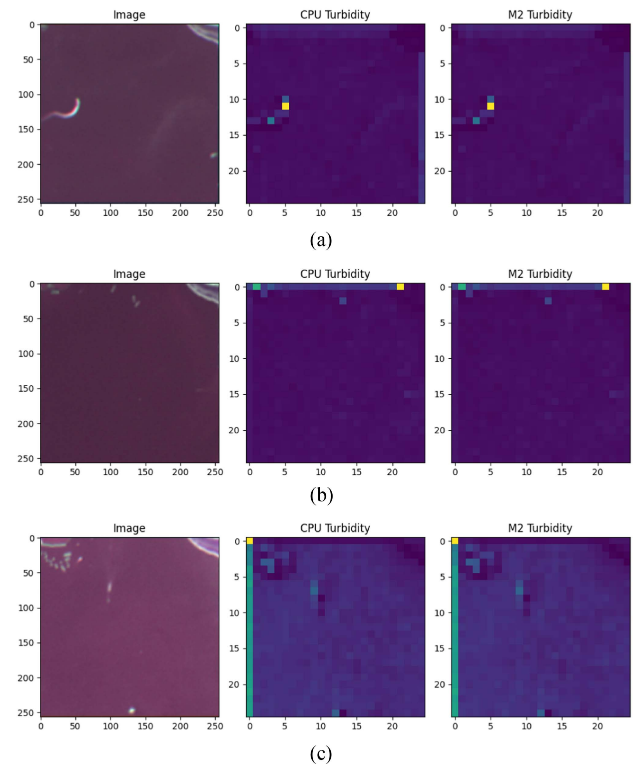


Fig. 9. Results for final AI onboard-based application confirming the feasibility through Myriad 2 implementation of the proposed model.

TABLE VIII
COMPARISON OF PH METRICS WITH OTHER STUDIES REVEALS THE SUPERIOR PERFORMANCE OF OUR ANALYSIS

pH	Methodology	RMSE	MAE
Our study	**AI4EDoET**	**0.1566**	**0.1191**
Kwong et al. [25]	ANN	0.172	0.147
Pereira et al. [49]	SMLR	0.85	–
Pereira et al. [49]	GP model	0.55	–

([25], [48] and [49]), to show the better performances of our method.

Making comparisons with pH parameters comparable to the case study proposed in our work has proven challenging from our knowledge. It has been difficult to find an article that addresses a study comparable to ours. This once again highlights the potential of our model and our application, which we propose as an onboard system alert for the upcoming Φsat-2 mission. It is also important to point out that as regards the feasibility of the model, our model has a size of approximately 4 MB, measured after the conversion to FP 16 used for the Myriad 2, and therefore we are well below the constraint set by ESA of 250 MB. Power constraint is fundamental for onboard application. In this case, no power constraint for the selected hardaware was specifically provided by ESA. Indeed, as discussed in the work by Furano et al. [21], the inferences on Myriad 2 of different models (models with sizes much larger than the one we considered) all result within a 1.2-W power envelope. Because of the high energy efficiency of the selected hardware and limited impact of the choice of the model on power results, no power test was performed.

B. Results With Myriad 2

As this operation is intended to be conducted on board Φsat-2, to demonstrate the functionality of the inference part on board Φsat-2, we present in this section some examples of images (in Fig. 9) computed as the output of our CNN on our CPU and the output calculated through the implementation on Myriad 2. The implementation on Myriad 2 involved conversion first to ONNX and then through Rupia, with rigorous testing throughout the process. From these qualitative results, it is clear that we could reproduce the same outputs obtained on the CPU also through Myriad 2; therefore the feasibility of the onboard model was successfully verified. Indeed, in support of the feasibility and compactness of our model, the inference time is 40.5 ms per inference, with a throughput of 24 Frame Per Second.

VII. Conclusion

The degradation of both water quality and quantity had been a pressing concern, with various human activities such as residential, agricultural, mining, industrial, power generation, and forestry operations having negatively impacted the aquatic environment. This phenomenon had been particularly pronounced in coastal areas, where marine contamination had manifested through a multitude of factors affecting the physical, chemical, and bacteriological properties of water, all of which had been linked to intended use and quality standards. The primary objective of this study has been to employ onboard AI techniques to monitor and detect contaminants in real-time since there are no methods available to do that on board satellite. This proposed approach represents a significant milestone in the context of real-time monitoring of water contaminants. As a result, it can facilitate the rapid generation of alerts and swift interventions when potentially hazardous events are on the horizon. In fact, the proposed approach, which involves processing data directly on board the satellite, allows for real-time responses and enhances the ability to manage environmental disasters. Among the main contributions of our work, it is worth underlining how our approach significantly diminishes response times for pollution detection. By performing the processing on board, anomalies can be identified and transmitted to the ground in real time. This eliminates the need to wait for the satellite data to be first sent to the ground and then processed, significantly reducing the overall response times. An example of this can be found in the study already introduced of Garcia et al. [22], which discusses the importance of onboard processing for improving response times in remote sensing applications.

The genesis of this endeavor has been traced back to the OrbitAI Challenge initiated by ESA Φ-lab, with the ultimate goal of deploying an AI-based application on board the Φsat-2 mission. There are several solutions that, stemming from this potential study on AI onboard techniques for contaminant monitoring, could be applied for advancements in future research. One such solution is represented by the employment of a different model such as the CatBoost Regressor. However, as illustrated in the previous sections, at present, integrating the Catboost Regressor into the satellite system is not feasible. From initial analyses, it is apparent that attention must be given not only to model goodness, assessed through metrics, but also to issues related to implementability on Myriad 2 for on board compatibility. Therefore, to ensure compatibility and successful deployment, we opted for the CNN framework, which aligns with the current technical requirements and available support tools. Yet, we are currently evaluating other regression models and conducting comparisons to determine if these models yield improved results. Therefore, we are proceeding with a series of analyses involving multiple ML and Deep Learning techniques. The objective is not only to extend the application to additional parameters, but also to identify new models capable of performing more accurate simulations. Another aspect may be related to expanding the dataset to enhance the network's generalization and yield more accurate results. Furthermore, given the network's capacity to generate valuable predictions for coastal monitoring, it could prove to be a significant asset in regions where acquiring these parameters is challenging. The transferability of the model, in fact, could be one of its major strengths, assisting in contaminant monitoring across various water zones. Another perspective development will entail broadening the monitoring scope to include the collection of chemical contaminants that directly impact human health, such as E. coli. In addition, future work could focus on comparing our model with others to identify the optimal solution for enhancing this study. In conclusion, this research not only provides valuable insights into coastal water quality monitoring, but also underscores the potential for cutting-edge technology and AI to make a meaningful contribution to safeguarding our precious coastal environments and ensuring the well-being of our Planet.

Acknowledgment

As highlighted in the text, the genesis of this research stems from the participation of the working group in the OrbitalAI Challenge, endorsed by ESA Φ-lab. The Team members would like to thank the ESA Φ-lab staff and the associated companies who followed them on this unbelievable pathway.

References

[1] B. Bisselink et al., "Impact of a changing climate, land use, and water usage on Europe's water resources: A model simulation study," EUR 29130 EN, Publications Office of the European Union, Luxembourg, 2018, doi: 10.2760/847068.

[2] M. Mainardis et al., "Wastewater fertigation in agriculture: Issues and opportunities for improved water management and circular economy," *Environ. Pollut.*, vol. 296, 2022, Art. no. 118755.

[3] U. Stein, G. Özerol, J. Tröltzsch, R. Landgrebe, A. Szendrenyi, and R. Vidaurre, "European drought and water scarcity policies," *Governance Drought Resilience: Land Water Drought Manage. Europe*, pp. 17–43, 2016.

[4] R. S. Araújo et al., "Water resource management: A comparative evaluation of Brazil, Rio de Janeiro, the European union, and Portugal," *Sci. Total Environ.*, vol. 511, pp. 815–828, 2015.

[5] S. Patra, S. Shilky, A. Kumar, and P. Saikia, "Impact of land use systems and climate change on water resource: Indian perspectives," in *Advances in Water Resource Planning and Sustainability*. Singapore: Springer Nature Singapore, 2023, pp. 97–110.

[6] T. Yao et al., "The imbalance of the Asian water tower," *Nature Rev. Earth Environ.*, vol. 3, no. 10, pp. 618–632, 2022.

[7] H. Alizadeh and V. Moshfeghi, "Spatial analysis of urban water vulnerability in cities vulnerable to climate change: A study in Ahvaz, Iran," *Int. J. Environ. Sci. Technol.*, vol. 20, no. 9, pp. 9587–9602, 2023.

[8] R. Lawford, A. Strauch, D. Toll, B. Fekete, and D. Cripe, "Earth observations for global water security," *Curr. Opin. Environ. Sustainability*, vol. 5, no. 6, pp. 633–643, 2013, aquatic and marine systems. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877343513001577

[9] World Health Organization and others, "Safer water, better health," 2019.

[10] W. Colglazier, "Sustainable development agenda: 2030," *Science*, vol. 349, no. 6252, pp. 1048–1050, 2015.
[11] M. Qasemi et al., "Characteristics, water quality index and human health risk from nitrate and fluoride in Kakhk city and its rural areas, Iran," *J. Food Comp. Anal.*, vol. 115, 2023, Art. no. 104870.
[12] M. Rahaman, M. Rahman, M. Mahbubur, and M. Nazimuzzaman, "Impact of salinity on infectious disease outbreaks: Experiences from the global coastal region," *Good Health and Well-Being*, pp. 415–424, 2020.
[13] M. P. Del Rosso et al., "On-board volcanic eruption detection through cnns and satellite multispectral imagery," *Remote Sens.*, vol. 13, no. 17, 2021, Art. no. 3479.
[14] R. Del Prete et al., "First results of vessel detection with onboard processing of Sentinel-2 raw data by deep learning," in *Proc. IGARSS 2023-2023 IEEE Int. Geosci. Remote Sens. Symp.*, 2023, pp. 6262–6265.
[15] G. Giuffrida et al., "Cloudscout: A deep neural network for on-board cloud detection on hyperspectral images," *Remote Sens.*, vol. 12, no. 14, 2020, Art. no. 2205.
[16] S. Lu et al., "Onboard AI for fire smoke detection using hyperspectral imagery: An emulation for the upcoming Kanyini Hyperscout-2 mission," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, pp. 9629–9640, 2024.
[17] T. Goudemant et al., "Onboard anomaly detection for marine environmental protection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, pp. 7918–7931, 2024.
[18] D. A. Maciel et al., "Evaluating the potential of CubeSats for remote sensing reflectance retrieval over inland waters," *Int. J. Remote Sens.*, vol. 41, no. 7, pp. 2807–2817, 2020. [Online]. Available: https://doi.org/10.1080/2150704X.2019.1697003
[19] P. D. Stasio, A. Sebastianelli, G. Meoni, and S. L. Ullo, "Early detection of volcanic eruption through artificial intelligence on board," in *2022 IEEE Int. Conf. Metrol. Extended Reality, Artif. Intell. Neural Eng.*, 2022, pp. 714–718.
[20] K. Thangavel, D. Spiller, R. Sabatini, P. Marzocca, and M. Esposito, "Near real-time wildfire management using distributed satellite system," *IEEE Geosci. Remote Sens. Lett.*, vol. 20, 2023, Art. no. 5500705.
[21] G. Furano et al., "Towards the use of artificial intelligence on the edge in space systems: Challenges and opportunities," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 35, no. 12, pp. 44–56, Dec. 2020.
[22] L. P. Garcia, "Advancements in on-board processing of synthetic aperture radar (SAR) data: Enhancing efficiency and real-time capabilities," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, early access, Jun. 03, 2024, doi: 10.1109/JSTARS.2024.3406155.
[23] F. Mohseni et al., "Ocean water quality monitoring using remote sensing techniques: A review," *Mar. Environ. Res.*, vol. 180, 2022, Art. no. 105701. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0141113622001465
[24] G. E. Adjovu, H. Stephen, D. James, and S. Ahmad, "Overview of the application of remote sensing in effective monitoring of water quality parameters," *Remote Sens.*, vol. 15, no. 7, 2023, Art. no. 1938. [Online]. Available: https://www.mdpi.com/2072-4292/15/7/1938
[25] I. Kwong, F. Wong, and T. Fung, "Automatic mapping and monitoring of marine water quality parameters in Hong Kong using Sentinel-2 image time-series and google earth engine cloud computing," *Front. Mar. Sci.*, vol. 9, 2022, Art. no. 871470.
[26] T.-N. Do et al., "Assessing surface water pollution in Hanoi, Vietnam, using remote sensing and machine learning algorithms," *Environ. Sci. Pollut. Res.*, vol. 30, no. 34, pp. 82230–82247, 2023.
[27] M. H. Rahaman, Roshani, M. Masroor, and H. Sajjad, "Integrating remote sensing derived indices and machine learning algorithms for precise extraction of small surface water bodies in the lower Thoubal river watershed, India," *J. Cleaner Prod.*, vol. 422, 2023, Art. no. 138563. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S095965262302721X
[28] M. Nazeer et al., "Evaluation of empirical and machine learning algorithms for estimation of coastal water quality parameters," *ISPRS Int. J. Geo- Inf.*, vol. 6, no. 11, 2017, Art. no. 360.
[29] T. Vakili and J. Amanollahi, "Determination of optically inactive water quality variables using Landsat 8 data: A case study in geshlagh reservoir affected by agricultural land use," *J. Cleaner Prod.*, vol. 247, 2020, Art. no. 119134. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0959652619340041
[30] E. Sharaf El Din, Y. Zhang, and A. Suliman, "Mapping concentrations of surface water quality parameters using a novel remote sensing and artificial intelligence framework," *Int. J. Remote Sens.*, vol. 38, no. 4, pp. 1023–1042, 2017.
[31] In-Situ Inc., "What are typical turbidity values in natural environments?," Accessed: Sep. 2024. [Online]. Available: https://in-situ.com/us/faq/water-quality-information/turbidity-faqs/what-are-typical-turbidity-values-in-natural-environments
[32] U. S. Environmental Protection Agency, "Water quality parameters factsheets," Accessed: Sep. 2024. [Online]. Available: https://www.epa.gov/awma/factsheets-water-quality-parameters
[33] J. Suzuki, M. Imamura, D. Nakano, R. Yamamoto, and M. Fujita, "Effects of water turbidity and different temperatures on oxidative stress in caddisfly (stenopsyche marmorata) larvae," *Sci. Total Environ.*, vol. 630, pp. 1078–1085, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0048969718306867
[34] R. Muoio, C. Caretti, L. Rossi, D. Santianni, and C. Lubello, "Water safety plans and risk assessment: A novel procedure applied to treated water turbidity and gastrointestinal diseases," *Int. J. Hyg. Environ. Health*, vol. 223, no. 1, pp. 281–288, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1438463918307363
[35] World Health Organization, *Guidelines for Drinking-water Quality*, 4th ed. Geneva, Switzerland: World Health Organization, 2011, Accessed: Aug. 05, 2024. [Online]. Available: https://www.who.int/publications/i/item/9789241549950
[36] F. Environmental, "ph - water quality parameters," n.d., Accessed: Aug. 09, 2024. [Online]. Available: https://www.fondriest.com/environmental-measurements/parameters/water-quality/ph/
[37] J. Jang et al., "Environmental Escherichia coli: Ecology and public health implications–a review," *J. Appl. Microbiol.*, vol. 123, no. 3, pp. 570–581, 2017.
[38] J. Jang et al., "Environmental Escherichia coli: Ecology and public health implications–a review," *J. Appl. Microbiol.*, vol. 123, no. 3, pp. 570–581, 2017.
[39] AI4EO Consortium, "PhiSat-2 mission overview," 2023. [Online]. Available: https://ai4eo.eu/wp-content/uploads/2023/02/Phisat-2_Mission_Overview_Web.pdf
[40] ESA-AI4EO, "phisat-2-simulator.ipynb," 2024. [Online]. Available: https://github.com/AI4EO/orbitalAI/blob/main/phisat-2/phisat-2-simulator.ipynb
[41] D. Moloney, B. Barry, R. Richmond, F. Connor, C. Brick, and D. Donohoe, "Myriad 2: Eye of the computational vision storm," in *2014 IEEE Hot Chips 26 Symp.*, 2014, pp. 1–18.
[42] E. Rapuano et al., "An fpga-based hardware accelerator for CNNs inference on board satellites: Benchmarking with myriad 2-based solution for the cloudscout case study," *Remote Sens.*, vol. 13, no. 8, 2021, Art. no. 1518. [Online]. Available: https://www.mdpi.com/2072-4292/13/8/1518
[43] V. Leon et al., "Improving performance-power-programmability in space avionics with edge devices: VBN on Myriad2 SoC," *ACM Trans. Embedded Comput. Syst.*, vol. 20, no. 3, pp. 1–23, 2021.
[44] A. Kyriakos et al., "Design and performance comparison of CNN accelerators based on the intel movidius Myriad2 SoC and FPGA embedded prototype," in *2019 Int. Conf. Control, Artif. Intell., Robot. Optim.*, 2019, pp. 142–147.
[45] C. Marantos et al., "Efficient support vector machines implementation on Intel/Movidius Myriad 2," in *2018 7th Int. Conf. Modern Circuits Syst. Technol.*, 2018, pp. 1–4.
[46] G. Meoni et al., "The OPS-SAT case: A data-centric competition for onboard satellite image classification," *Astrodynamics*, pp. 1–22, 2024.
[47] F. Razzano et al., "Monitoring water contaminants in coastal areas through ML algorithms leveraging atmospherically corrected Sentinel-2 data," in *Proc. IGARSS 2024-2024 IEEE Int. Geosci. Remote Sens. Symp.*, 2024, pp. 1046–1050. [Online]. Available: https://arxiv.org/abs/2401.03792
[48] S. Hafeez et al., "Comparison of machine learning algorithms for retrieval of water quality indicators in Case-II waters: A case study of Hong Kong," *Remote Sens.*, vol. 11, no. 6, 2019, Art. no. 617. [Online]. Available: https://www.mdpi.com/2072-4292/11/6/617
[49] O. J. R. Pereira et al., "Estimating water pH using cloud-based landsat images for a new classification of the Nhecolândia lakes (Brazilian pantanal)," *Remote Sens.*, vol. 12, no. 7, 2020, Art. no. 1090. [Online]. Available: https://www.mdpi.com/2072-4292/12/7/1090
[50] J. R. Koza, "Evolving a computer program to generate random numbers using the genetic programming paradigm," in *Proc. ICGA*, 1991, pp. 37–44.
[51] G. Mateo-García et al., "Onboard cloud detection and atmospheric correction with deep learning emulators," in *Proc. IGARSS 2023-2023 IEEE Int. Geosci. Remote Sens. Symp.*, 2023, pp. 1875–1878.

Francesca Razzano (Student Member, IEEE) received the graduation degree in electronic engineering for automation and telecommunications from the University of Sannio, Benevento, Italy, in 2023. She is currently working toward the Ph.D. degree in information and communication technology and engineering with the University of Parthenope, Naples, Italy, under Supervisor Prof. Gilda Schirinzi and Co-Supervisor, Prof. Silvia L. Ullo.

She has authored or coauthored two papers in reputed conferences for the sector of Remote Sensing. Her research interests include Remote Sensing and Satellite data analysis, Artificial Intelligence techniques for Earth Observation, and data fusion with a particular focus in water quality monitoring.

Dr. Razzano is affliated with the Geoscience and Remote Sensing Society (GRSS) and Aerospace and Electronic Systems Society (AESS). She is a Member of a Team competing for OrbitalAI Challenge by ESA-Φ-lab arriving to the last five finalist-Team.

Pietro Di Stasio (Member, IEEE) received the master's degree in electronics engineering for automation and telecommunications from the University of Sannio, Benevento, Italy, in 2024.

He was a Visiting Researcher with Φ-lab, European Space Agency (ESA), Frascati, Italy. His research interests include Earth Observation, Remote Sensing, and the application of AI onboard for satellite data analysis.

Dr. Stasio was the recipient of a research scholarship in 2022 at University of Sannio, under Supervisor Prof. Silvia L. Ullo, following the Bachelor's degree. His work during this scholarship focused on utilizing AI onboard for volcanic eruption detection. He is affliated to the Geoscience and Remote Sensing Society (GRSS) and Aerospace and Electronic Systems Society (AESS). He is a Member of a Team competing for OrbitalAI Challenge by ESA-Φ-lab arriving to the last 5 finalist-Team.

Francesco Mauro (Student Member, IEEE) received the master's degree (with laude and career mention) in electronic engineering for automation and telecommunications, in 2022, from the University of Sannio, Benevento, Italy, where he is currently working toward the Ph.D. degree in researching the emerging area of Quantum Machine Learning for Earth Observation (QML4EO).

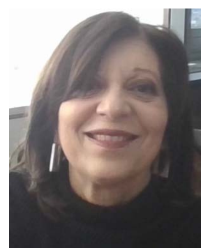
He was a Visiting Researcher with the Φ-lab, European Space Agency, Frascati, Italy, and still collaborates with the Φ-lab on topics related to Quantum Machine Learning applied to Earth observation. He has authored or coauthored several papers for the sector of remote sensing and has presented his work in reputed international conferences. His research interests include remote sensing and satellite data analysis, artificial intelligence, and quantum machine learning techniques for Earth observation.

Gabriele Meoni (Member, IEEE), received the Laurea degree in electronic engineering and the Ph.D. degree in information engineering from the University of Pisa, Pisa, PI, Italy, in 2016 and 2020, respectively.

During the doctoral studies, he developed skills in digital and embedded systems design, digital signal processing, and artificial intelligence. He is currently an Assistant Professor with the Delft University of Technology, Delft, Netherlands. Since 2020, he has been a Research Fellow with the ESA Advanced Concepts Team. His research interests include machine learning, embedded systems, and edge computing.

Marco Esposito (Member, IEEE) received the Ph.D. degree in aerospace engineering, development of airborne remote sensing systems and data interpretation for environmental monitoring from the University of Naples, Naples, Italy.

Having worked on various aerospace projects in Naples, most recently as a Coordinator with the Environmental Research Aircraft Laboratory, Italian National Research Council, Rome, Italy, he moved to cosine Measurement Systems BV, Warmond, Netherlands. Having gained experience as a System and Performance Engineer, a Project Manager, a Program Manager, and a Team Leader with cosine Measurement Systems BV, he became the Managing Director of cosine as well as of cosine Remote Sensing B.V. He has been leading a variety of research and development activities with cosine Measurement Systems, all focusing on the miniaturization of optical instruments for Earth Observation and Planetary Exploration, including NightPod, TropOMI, and Mobile Asteroid Surface Scout (MASCOT) on the Hayabusa 2 Sample Return Mission. He has led the development, demonstration in orbit, and commercialization of the HyperScout product series, the first ever miniaturized imager able to connect hyperspectral, thermal imaging, and artificial intelligence techniques in one compact product for the Earth Observation Market.



Gilda Schirinzi (Senior Member, IEEE) received the graduation degree (cum laude) in electronic engineering from the University of Naples Federico II, Naples, Italy, in 1983.

From 1985 to 1986, she was with European Space Agency, ESTEC, Noordwijk, The Netherlands. In 1988, she joined the Istituto di Ricerca per l'Elettromagnetismo e i Componenti Elettronici (IRECECNR), Naples, Italy. In 1998, she joined the University of Cassino, Cassino, Italy, as an Associate Professor of telecommunications, and in 2005, she became a Full Professor. Since 2008, she has been with the University of Naples "Parthenope," Naples, Italy. Her main research interests include the field of signal processing for remote sensing applications, with particular reference to synthetic aperture radar interferometry, and tomography.

Silvia Liberata Ullo (Senior Member, IEEE) received the graduation degree in electronic engineering from the University of Naples, Naples, Italy, in 1989, and the M.Sc. degree in management from the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 1992.

Since 2004, the author is a Researcher and Teacher with the University of Sannio, Benevento, Italy. The author is also President of IEEE AES Italy Chapter, Industry Liaison for IEEE Joint ComSoc/VTS Italy Chapter since 2018, National Referent for FIDAPA BPW Italy Science and Technology Task Force (2019–2021), Member of the Image Analysis and Data Fusion Technical Committee (IADF TC) of the IEEE Geoscience and Remote Sensing Society (GRSS) since 2020, and Member of Academic Senate and Ph.D. Professors' Board. Courses: Signal theory and elaboration, Telecommunication networks (Bachelor program); Earth monitoring and mission analysis Lab (Master program), Optical and radar remote sensing (Ph.D. program). The author has authored or coauthored more than 90 research papers, many book chapters, and served as Editor for two books. The author's research interests include signal processing, radar systems, sensor networks, smart grids, remote sensing, satellite data analysis, machine learning, and quantum ML.

Silvia Liberata Ullo is Associate Editor for relevant journals (*IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (IEEE JSTARS), *MDPI Remote Sensing*, *IET Image Processing*, *Springer Arabian Journal of Geosciences and Recent Advances in Computer Science and Communications*) and Guest Editor for many special issues.

Open Access funding provided by 'Università degli Studi di Napoli "Parthenope"' within the CRUI CARE Agreement