



## Integration of satellite surveillance and metagenomics for the monitoring and protection of water basins from oil spills

Emilio D'Ugo<sup>a,\*</sup>, Arghya Mukherjee<sup>b</sup>, Roberto Giuseppetti<sup>a</sup>, Matteo Tucci<sup>c</sup>, Paola Bucci<sup>a</sup>, Federico Aulenta<sup>c,d</sup>, Giovanni Laneve<sup>e</sup>, Fabio Magurano<sup>a</sup>

<sup>a</sup> Department of Infectious Diseases, Istituto Superiore di Sanità, Rome, Italy

<sup>b</sup> Department of Food Biosciences, Teagasc Food Research Centre, Fermoy, Co. Cork, Ireland

<sup>c</sup> Water Research Institute (IRSA) National Research Council (CNR), Monterotondo (Rome), Italy

<sup>d</sup> National Biodiversity Future Center, 90133 Palermo, Italy

<sup>e</sup> Scuola di Ingegneria Aerospaziale, "Sapienza" University of Rome, 00138 Rome, Italy

### ARTICLE INFO

#### Keywords:

Remote sensing, satellite imagery  
Oil spill  
Metagenomics  
Water basins  
Freshwaters  
Petroleum hydrocarbons  
Microbial ecology  
Dibenzothiophene (DBT)  
Desulfurization  
Lakes

### ABSTRACT

The planetary crisis regarding water resources means that new methods are needed to monitor large areas of water basins that are threatened by chemical and natural pollutants, together with climate change. With the aim to detect oil spill, we applied an algorithm, which could consistently and reliably confirm the presence of oil in four polluted lake waters analyzed. Combined algorithm application and metagenomic analysis from the spill areas that had been detected by the satellite identified drivers of the microbial response to oil pollution. Based on ortholog abundances, metabolic pathway reconstruction carried out in PICRUSt2 highlighted the degradative capacity of these microorganism. These microorganisms could be suitable candidates for treatment of crude oil, aromatic hydrocarbons and the desulfurization of persistent petroleum substances like dibenzothiophene.

Environmental changes have been analysed with the combination of satellite monitoring and metagenomic in other studies. Red snow phenomenon, in Franz Josef Land's Arctic observed by satellite, is the result of a microbial succession dominated by *Chlamydomonas nivalis*, a unicellular, red-colored photosynthetic green algae. Similarly, satellite monitoring and metagenomic monitoring were used to assess the impact on coral reefs of a huge quantity of mud spill from iron ore mining on the Abrolhos Bank reef in Brazil.

In our study, the combination of satellite sensing and metagenomics analyses offer useful tools for the real-time monitoring of water bodies threatened by oil spills, as well as for the design of recovery strategies based on the use of valuable hydrocarbonoclastic microorganisms.

### 1. Introduction

Freshwater contamination is a global concern. The enormous impact of the natural and anthropogenic organic substances that are constantly released into the environment, requires better knowledge of the chemical status of Earth's surface water (Sousa et al., 2018). The identification of chemical contaminants over large areas can provide indications regarding the type of danger and the need for timely interventions for the conservation of water supplies and its services. Emerging satellite technologies could allow the monitoring of sensitive water basins whose quality is important for human and animal health. Aqueduct management companies that draw drinking water from large water basins such as lakes, threatened by industrial plants, could make use of a

multispectral satellite surveillance system which, through validated algorithms, identifies the presence of pollutants. The Sentinel-2 satellite is part of the European Copernicus program and can capture various water quality parameters, such as chlorophyll concentration and turbidity levels, which are important indicators of the health and pollution levels of water bodies (Bramich et al., 2021, Drusch et al., 2012, Rodriguez-Benito et al., 2020). In a previous study it was studied the reflectance variation of the surface in drinkable waters of Lake Pertusillo during the presence of hydrocarbon due probably to oil spill from near oil plant (Kolokoussis and Karathanassi, 2018). Using Multispectral Instrument (MSI) of Sentinel 2, the oil presence was evaluated with a developed algorithm. Although, studies underlined the potential of Sentinel 2 to identify oil spills in oceans (Laneve et al., 2022), not

\* Corresponding author

E-mail address: [emilio.dugo@iss.it](mailto:emilio.dugo@iss.it) (E. D'Ugo).

<https://doi.org/10.1016/j.envadv.2024.100498>

Received 27 November 2023; Received in revised form 2 January 2024; Accepted 11 February 2024

Available online 13 February 2024

2666-7657/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

validated algorithm are available. Therefore, the validation of algorithms for pollutants such as oil, capable of inducing reflectance variations on the water surface, represents an objective for the development of multispectral and hyperspectral satellites.

Environmental microbial metagenomics (Andreote et al., 2012, Mukherjee et al., 2017, Yang et al., 2012) has been successful in obtaining complex information about microbial communities in aquatic ecosystems, including changes in microbial communities during pollution events (D'Ugo et al., 2021, Mason et al., 2012, Mukherjee et al., 2017) on the type of pollutant present (Sagova-Mareckova et al., 2021, Smith et al., 2015) and on microbial proxies for oil pollution (dos Santos et al., 2011), among others. The microbial response to petroleum hydrocarbon contamination in freshwater environments is not well-documented due to the rare occurrence of oil spills. However, information from metagenomic analysis can reveal the presence of microorganisms important for the protection of water basins. A combination of remote sensing technologies and metagenomic methods can be used for the large-scale spatial and temporal monitoring of polluted environments both at macro- and micro-levels.

In this study, using Sentinel-2, we validated and applied an algorithm, the Visible Near Red Index (VNRI), for the observation of hydrocarbon oil spills in five lakes located in Europe, Africa, and America. For one lake (Pertusillo, Italy), guided by the satellite maps generated by algorithmic processing that indicated the polluted areas, we studied the microbial community of the affected areas. In Lake Pertusillo, polluted areas demonstrate changes in the lake microbial community due to hydrocarbon pollution, with hydrocarbonoclastic microbial communities and functional pathways involved in hydrocarbon degradation.

Such combinations may be important in the development of powerful monitoring technologies for the rapid detection of oil spills in watersheds and help improve the management of drinking water supplies also through the planning of appropriate microbial-based restoration measures.

## 2. Materials and methods

### 2.1. Study site and sample collection

Samples were collected from Lake Pertusillo (Supporting Information, SI, Fig. S1) after 400 tons of crude oil were accidentally spilled out in February 2017 from the oil extraction facility. Total Petroleum Hydrocarbons (TPH) determination performed by civic associations (EPA 3546A 2007 and 8270D Methods) revealed the presence of oil in the lake (<https://covacontro.org/idrocarburi-nel-pertusillo-presenti-in-tre-campioni-su-tre-oltre-a-fosfati-e-manganese/>). The same samples used for TPH determination were splitted for metagenomic analyses: Plume-UP and Plume-SP were collected from the water surface, whereas the Plume-DW sample was collected below the surface (0.2m) (Table 1). Samples Plume-UP, Plume-SP and Plume-DW were collected with 1-L Pyrex glass bottles and stored in ice (4°C) boxes. Water samples were later processed using a previously described polysulfone-based ultrafiltration protocol (D'Ugo et al., 2016). The elutriate (50 mL) was aliquoted and stored at -80°C and used for subsequent laboratory investigations.

### 2.2. Spectral acquisition of hydrocarbons

To develop a methodology to establish reflectance in water basins, that are characterized by the presence of petroleum, a spectral acquisition campaign funded by PDVSA (Petróleos de Venezuela, S.A), was conducted in 2011 on Lake Maracaibo (Supporting Information, SI). Lake Maracaibo is one of the world's richest and most centrally located petroleum-producing regions. The field campaign included instruments such as a FieldSpec ASD (a portable spectroradiometer that operates between 0.350 and 2.5 µm), a USB 2000 (a highly portable spectrometer that operates in the region between 0.200 and 1.100 µm), a Micro FT-IR spectrometer (a portable instrument that measures the radiation emitted by a body in the 2 to 16 µm part of the electromagnetic spectrum), and a FLIR SC660 (a camera that operates in the thermal infrared region [7.5 -

**Table 1**  
Satellite and microbiome survey of contaminated and uncontaminated lakes.

Sample ID	Collection Site	Sampling depth	SRA Run ID	Total Hydrocarbon content µg/L	Sampling Date	Sampling/Satellite survey Coordinates	Reflectance based index values (VNRI)	Satellite survey of areas of the lake impacted by oil spills (km <sup>2</sup> ). The estimate is based on FOC
Plume_UP	Masseria Crisci point at Pertusillo lake	Surface	SRX5531070	192	February 27, 2017	40.28977 N 15.95180 E	0.161	0.315 corresponding to the 5.5% of the whole lake surface
Plume_DW	Masseria Crisci point at Pertusillo lake	20 cm	SRX5531072	192	February 27, 2017	40.28977 N 15.95180 E	0.161	
Plume_SP	Spartifave 2 at Pertusillo lake	Surface	SRX15224861	213	February 27, 2017	40.29687 N 15.93075 E	0.172	
LD	Pertusillo lake damming	NA*	NA	87	August 3, 2017	40.27522 N 15.99157 E	0	0
WTZA	Wietze pond	Surface	SRR12527656(WT22)	Asphalt lake	November 5, 2019	52.666582 N 9.819021 E	NA	NA
			SRR12527655(WT23)	Asphalt lake	November 5, 2019	52.666597 N 9.819031 E	NA	NA
			SRR12527657(WT21)	Asphalt lake	November 5, 2019	52.666600 N 9.819015 E	NA	NA
NA	Al Borma lake	NA	NA	NA	March 16, 2023	31.715 N, 9.182 E	0.206	0.2095 corresponding to the 59 % of the whole lake surface
NA	Zenaiga lake	NA	NA	NA	March 16, 2023	31.655 N, 9.167 E	0.155	0.004 corresponding to the 9 % of the whole lake surface
MG	Madonna Grumentina Pertusillo lake	20 cm	SRR6255875	NA Uncontaminated site during the sampling	January 15, 2016	40.29 N, 15.92 E	0	0
ALB_19.1	Lake Alboraj	1 metre	SRR14718350	Pristine lake	October, 2019	38°35' N, 1°39' W	NA	NA
ALB_19.4	Lake Alboraj	4 metres	SRR14718349	Pristine lake	June 25, 2019	38°35' N, 1°39' W	NA	NA

NA\* not available.

13.0  $\mu\text{m}$ ]. The image consists of  $640 \times 480$  pixels).

### 2.3. Sentinel-2A/2B satellite monitoring by algorithm

Researchers commonly consider radar images more suitable to detect oil spills even if several sources of false alarm exist, in particular in the case of inner water bodies (lakes) like, the absence of wind near the coasts, the presence of floating vegetation (e.g. *Lemna* sp.). Indices in the VIS/NIR region of the electromagnetic spectrum has been mainly developed to detect classical water pollutant as chlorophyll, TSS (Total Suspended Solid), CDOM (Colored Dissolved Organic Matter), turbidity and trophic state index. Very few studies have been devoted at detecting oil spill in lakes using both radar and VIS/NIR images. Based on the experience gained by monitoring the spectral behaviour of the-oil polluted water of Lake Maracaibo, we developed our own index described in Laneve et al. 2022. The VNRI relates the reflectance values of visible green and red bands (0.560  $\mu\text{m}$ : MSI channel 3; 0.665  $\mu\text{m}$ : MSI channel 4) and the near red band (0.740  $\mu\text{m}$ : MSI channel 6). The equation highlights the importance of reflectance values in the band near the red (0.740  $\mu\text{m}$ : MSI channel 6) considered specific for the presence of oil, therefore denser oil-polluted waters correspond to higher positive values.

The equation (Eq. (1)) is as follows:

$$VNRI = -\frac{2 * r_{560} - r_{665} - r_{740}}{r_{560} + r_{665} + r_{740}} \quad (1)$$

The VNRI index has negative values, for reflectances that are prevalent in the visible spectrum, whose peaks overlap in clean water and in oily water (channels 0.560  $\mu\text{m}$  and 0.665  $\mu\text{m}$ ), as in the cases shown in the insert in Fig. 1. The index assumes positive values in the cases in which the reflectance close to red prevails, which is a typical characteristic of oily waters (channel 740) (see insert Fig. 1A). For the present study, we computed VNRI indices on Level 2 images, which were atmospherically corrected, and we acquired images between June 2016 and May 2021. A detailed description of Sentinel Satellites is provided in SI. An area historically known as petroleum contaminated site, Wietze pond in (Germany) was used as a positive control for 16S rRNA based investigations (Table 1, Bioproject PRJNA612814, Eze et al., 2021).

### 2.4. DNA extraction and sequencing and 16S rRNA amplicon sequence analysis

Each collected water samples were centrifuged at 6000 rcf for 10 min at room temperature. The pellet of 5 milliliters, was digested with lysozyme (Sigma) and the total environmental DNA was extracted using the Genomic DNA tissue kit (Machery-Nagel, Thermo Fisher, Italy). Extracted environmental DNA was quantified spectrophotometrically (NanoDrop, Thermo Fisher) and fluorometrically (Qubit, Thermo Fisher, Italy). The V3-V4 hypervariable region of the prokaryotic ribosomal small subunit RNA (16SrRNA) was amplified from total environmental DNA with primers 341F (5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG-3') and 785R (5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC-3').

16S rRNA amplicon sequencing (550 bp) was then performed by the Bio-Fab research Laboratory (Rome, Italy) using the Illumina MiSeq sequencing platform with the  $2 \times 300$  bp read modus (v3 chemistry, Illumina). All sequencing data generated in the study has been deposited in the NCBI Sequence Read Archive (SRA) under BioProject number PRJNA412797 (Table 1; Plume\_SP: SRX15224861; Plume\_UP: SRX5531070; Plume\_DW SRX5531072).

The 16S rRNA datasets used in this study are outlined in Table 1. Further information about each collected sample is reported in SI.

### 2.5. Metagenome reconstruction

Quality and contaminant-filtered ASVs from rarefied datasets were used as inputs for the prediction of metagenomes from 16S rRNA sequences by PICRUSt2 (v2.5.0) (Douglas et al., 2020). Metagenomes were predicted in terms of both KEGG ortholog and Enzyme Commission (EC) number abundances. For each sample, the weighted nearest sequence taxon index (NSTI) value was calculated during metagenome prediction. Based on the results, the Weitz samples were removed from further PICRUSt2 analysis (Table S2). Briefly, NSTI values allow an estimation of prediction accuracy for PICRUSt2 where a NSTI value of above 0.15 is deemed to be unreliable and is strongly recommended to be removed from the analysis (<https://github.com/picrust/picrust2>). In addition to predicted metagenomes, taxonomic contributions to each predicted ortholog were also computed. Subsequently, the taxa that contribute to KEGG orthologs and hydrocarbonoclastic genes of interest were determined at genus level and graphed. Based on ortholog abundances, metabolic pathway reconstruction was also carried out in PICRUSt2. To this end, the MetaCyc metabolic pathway database was used to define and enumerate the pathways in each predicted metagenome (Caspi et al., 2020). All visualisations were carried out in R, using ggplot2 (Wickham et al., 2016).

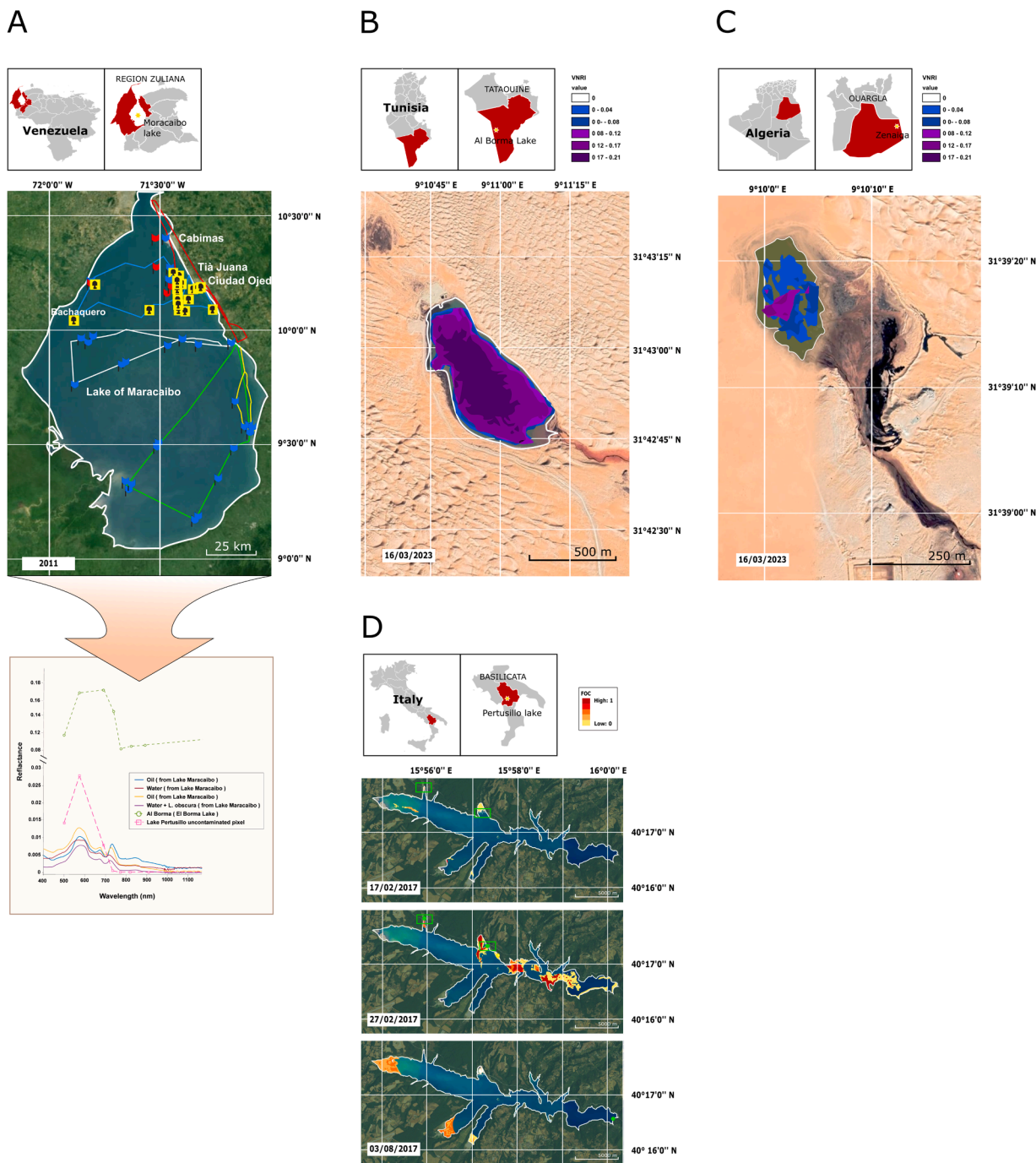
## 3. Results

### 3.1. Infield and satellite reflectance of lakes with the presence of hydrocarbon spills

To study the spectral behaviour of the oil that floats on the surface of lakes, a field study was conducted in Lake Maracaibo, which has been subjected to repeated oil spills ([www.washingtonpost.com/world/2021/10/07/oil-pollution-lake-maracaibo-venezuela](http://www.washingtonpost.com/world/2021/10/07/oil-pollution-lake-maracaibo-venezuela)). The measurement campaign was carried out from during February to March 2011 with field instruments and then the channels of the electromagnetic spectrum (EMS) were defined in which it is easier to distinguish the oil contaminated water surface. The area covered by the campaign is shown in Fig. 1A. The set of instruments used was selected to cover the whole optical region (0.35 - 14  $\mu\text{m}$ , see materials and methods) of the EMS that is usually covered also by satellite sensors. The Fig. 1A shows the continuous spectral trends comparing the Maracaibo water of the different sections monitored using boats or observed by helicopter. Except for the chlorophyll peak at 705 nm, the surface waters of the oil polluted Lake Maracaibo appear to have overlapped spectral peaks in the visible range of the EMS (500-700 nm). This trend is the typical reflectance for eutrophic water at 560 and 705 nm. At slightly higher wavelengths, a gap begins to appear that highlights an area of differentiation between clean and oily water (inset Fig. 1 A). The reflectance spectra shown in Fig. 1A represent average values taken in different sites during the cruise carried out on Lake Maracaibo in 2011. These observations were confirmed with the multispectral instrument (MSI) present on Sentinel-2 by analyzing the other lakes selected for the study (Table 1) that were also subject to oil spills. The insert of Fig. 1 A shows the reflectance trend (green line) for Al Borma in the presence of heavy oil spill, and the reflectance trend (red dotted line) for Lake Pertusillo (Supporting Information: SI, Fig.S1) in the absence of oil. The results of the observation of Sentinel-2 satellite MSI, which acquires the solar radiation reflected between 440 and 2200 nm divided into 13 multi-spectral channels, highlight characteristics that can be superimposed on the readings with in field spectral instruments of Lake Maracaibo (Fig.1A).

### 3.2. Sentinel-2- visible near red index (VNRI) Lake images confirm oil presence in affected areas

To date, no specialized indices have been developed to map oil spills. Considering the optical behavior of oil-water dispersion in natural



**Fig. 1.** Infield and satellite observation. A. Map of the sites of Lake Maracaibo where the spectral and radiometric measurements were performed. The yellow symbols represent the points where thermal maps were collected from a helicopter. Inset: examples of oil spectra and water spectra collected at Lake Maracaibo in 2011. Note the characteristic peak of reflectance at 0.705 micron, due to *Lemna obscura*, a typical floating vegetation of Lake Maracaibo. Reflectance Spectra collected *in situ* are compared with two spectra extracted from Sentinel-2 images of Al Borma (16th March 2023) and Lake Pertusillo (3rd August 2017). In Al Borma case, the spectrum refers to a pixel contaminated by oil; in Pertusillo's case, the pixel is not contaminated by oil. B. VNRI map of the polluted deserts Lake Al Borma (Tunisia) close to an oil field. C. NRI map of a polluted deserts lake near Zenaiga (Algeria) and close to an oil field. D. Sentinel-2 surveillance of Lake Pertusillo. Maps of FOC (fraction of pixel covered) obtained as described in the text, based on Sentinel-2 images superimposed onto a Google map of Lake Pertusillo. Upper Panel, FOC map of 17th February, 2017 of Lake Pertusillo. The green boxes represent the two supposed point pollution areas discussed in the text. The Orange/brown plume on the left side of the lake corresponds to areas where the presence of oil starts to appear. In accordance with the polygon which, during winter, delimits (in blue) the left side of the lake (Agri River mouth) is covered by water. Middle Panel, green points with green boxes represent the point in which the ground data were collected on 27 of February 2017 (see Table 1). Orange/brown/violet correspond to the presence of oil in the lake. Below Panel, Map of FOC obtained on 3rd August, 2017 of Lake Pertusillo. Green point represents the site where data on oil presence was collected on the same day (see also Table 1). No evidence of the presence of oil is given by the VNRI index. In this case orange/brown areas (left side of the lake) correspond to areas where in summer water scarcity makes the lake bottom visible.



waters (see insert Fig. 1A) and the results of our previous studies, we used a reflectance ratio index (Visible Near Red Index) described in Laneve et al., 2022. The VNRI map gives an estimate of the amount of oil in a pixel of the satellite image. Assuming as completely covered by oil a pixel where the VNRI index reaches its the maximum value we can try to build a FOC (Fraction of Oil) map, that could allow to estimate the total area covered by the oil spill. For the validation of the VNRI index, we analysed water basins whose frequent contamination or oil presence is known such as in the desertic lakes Al Borma on the border between Tunisia and Algeria, Zenaiga (Algeria), and Lake Pertusillo (Italy). Lake Pertusillo and the Wietze pond were selected in order to access chemical data and databases on microbial communities (Table 1).

The region of the electromagnetic spectrum considered for oil spill detection covers the VIS-NIR bands. Taking into account the continuous trend observed in the oiled and non-oiled waters of Maracaibo, the index was based on the reflectance measured at 0.740  $\mu\text{m}$  from channel B6, on board Sentinel-2. In this region the channels are not affected by the chlorophyll reflectance peaks at 0.705  $\mu\text{m}$  (B5).

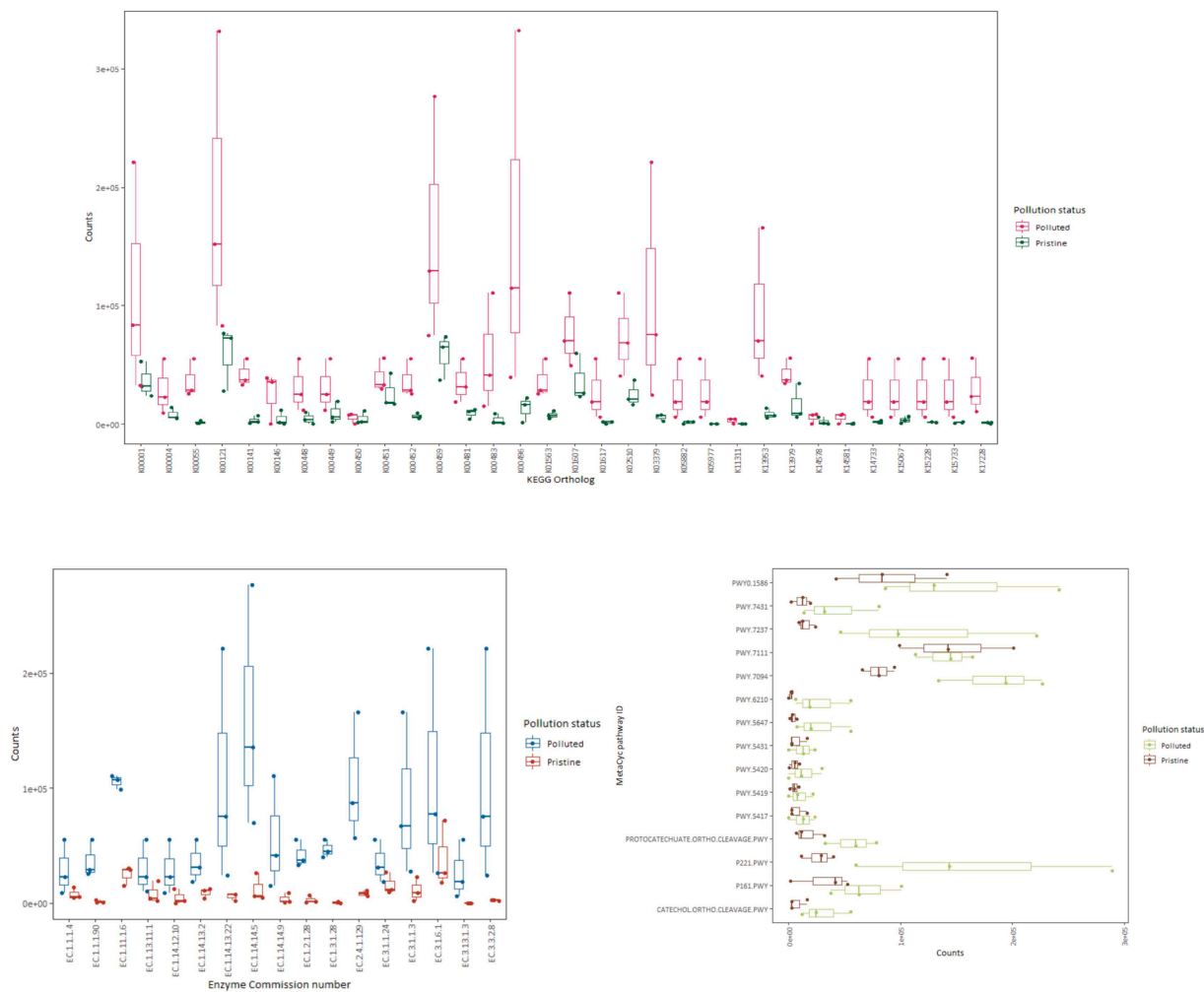
The VNRI analysis successfully produced a map of the desert water bodies near the rich oil fields of Al Borma and Zenaiga that were contaminated (Figs. 1B-C). Based on the map of VNRI, we were able to try and estimate the extension of the area affected by the oil spill (see Table 1). This was achieved by considering the pixels for which the VNRI reaches its maximum to be completely covered by oil. Consequently, the other pixels will be covered by oil in the ratio VNRI/VNRI<sub>max</sub>. We named this ratio FOC (Fraction of Cover). Therefore, the FOC values range from 0 to 1. The images acquired on March 16, 2023 (image S2B\_MSI-L2A\_20230316T100029\_N0509\_R122\_T32SNA) show a substantial percentage of oil present in Lake Al Borma close to an oil extraction plant (<https://meshkal.org/the-cost-of-oil-in-tunisia-desert/>). Similar observations were obtained in Lake Zenaiga where VNRI indexed a lower percent of oil covering the water basin (see FOC values Table 1 and Fig. 1C). In Fig. 1D, the FOC (VNRI/VNRI<sub>max</sub>) approach was applied to three images of Lake Pertusillo obtained on 17 February 2017, 27 February 2017 and 3 August 2017. The red points represent the area where the VNRI index reaches its maximum value, then maximum concentration of oil in the pixel. Therefore, according to our approach in such pixels the FOC = 1. During this time, the fraction area covered by oil (FOC, represented in yellow-red) grew from 0.086 km<sup>2</sup> (17 February) to 0.315 km<sup>2</sup> (27 February) and then went back to almost 0 (3 August). FOC maps probably provide more understandable evidence of the impact of the phenomenon compared to only VNRI. In Lake Pertusillo after a long cloudy period, on February 17, 2017, FOC increased only very locally in the northern and southern tributary regions and along a suspect dark flow mark across the western basin (Fig. 1D, upper panel). On this day, FOC index areas (expressed in square kilometres, Table 1) show the presence of hydrocarbons with an area of 0.086 km<sup>2</sup> affected by the oil spill. This extends in the following days (until February 27, 2017) to reach 0.315 km<sup>2</sup> (Fig. 1E middle panel). On February 27, 2017, the VNRI in the west of Lake Pertusillo remained below the maximum VNRI index value observed (0.42) near the Agri River inflow in the west, but there was a significant vortex with maximum VNRI observed from the lake's centre toward the east. The extension of this area on this day included the sites of Masseria Crisci and Spartifave, from where the water samples shown in Table 1 were collected. Interestingly, these sites already highlighted significant FOC values on February 17, 2017 (Fig. 1 upper panel). This could indicate a source of point pollution from where some of the spilled oil could have been released into the lake.

### 3.3. Lacustrine microbiomes from oil-polluted areas detected by VNRI index reveal hydrocarbonoclastic patterns

One of the major challenges of this study was to study the microbial succession in contaminated areas measured with the VNRI that showed changes over time. This may indicate a state of pollution through microbial indicators. To understand the microbial community structures of

Lake Pertusillo during the oil spill, 16S rRNA gene sequence analysis of lake samples was performed before (Madonna Grumentina, MG) and during the oil spill in polluted areas detected by the VNRI index (Masseria Crisci: PLUME-DOWN and PLUME-UP; and Spartifave: PLUME-SP) (Table 1). 16S rRNA datasets from two pristine lake water samples from Lake Alboraj (ALB\_19\_1 and ALB\_19\_4) and three chronically hydrocarbon-contaminated lake water samples (WTZ1, WTZ2, WTZ3) from the small asphalt pond of Wietze in Germany were used as references in relation to hydrocarbon contamination (Fig. S2). The polluted samples show recurrence of hydrocarbonoclastic genera and reduced biodiversity (SI, Fig. S3). Metagenome prediction and hydrocarbonoclastic potential of lacustrine microbiomes indicates a strong hydrocarbonoclastic community in Lake Pertusillo (Fig. 2).

To functionally characterise lake microbiomes before and during oil spills, protein ortholog prediction and metabolic pathway reconstruction were carried out with PICRUSt2 (Fig. 2A-B). To this end, the MetaCyc metabolic pathway database was used to define and enumerate the pathways in each predicted metagenome (Fig. 3C). The asphalt pond samples of Wietze were removed from further PICRUSt2 analysis due to unsuitable NSTI values (see methods section, Table S3). Ortholog prediction was carried out for both KEGG orthologs and Enzyme Commission annotations. The comparison of counts for hydrocarbonoclastic genes and pathways between normalised polluted samples (Pertusillo lake after the spill) and pristine samples (Pertusillo Lake before the spill and Lake Alboraj) revealed the hydrocarbonoclastic metabolic potential of the lake Pertusillo microbiome during an oil spill (PLUME\_DOWN, PLUME\_UP, and PLUME\_SP), where multiple metabolic pathways involved in the degradation of aliphatic and aromatic hydrocarbons were detected (Fig. 2C, Table S2). Significant differences in hydrocarbonoclastic gene counts were observed between the contaminated and uncontaminated samples, particularly the uncontaminated Pertusillo (MG) sample and contaminated ones (Fig. 2A-B, Table S2). The hydrocarbonoclastic orthologs that were identified are potentially capable of oxidizing aliphatic and aromatic hydrocarbons and the various alcoholic intermediates, aldehydes and ketones up to the introduction of catabolites in the beta oxidation of fatty acids and in the cycle of tricarboxylic acids. These include alkane 1-monooxygenase (K00496, EC: 1.14.15.3); haloalkane dehalogenase (K01563, EC: 3.8.1.5), nitronate monooxygenase (K00459, EC: 1.13.12.16) and Alkanesulfonate monooxygenase (EC: 1.14.14.5) (Fig. 2A-B). Aliphatic hydrocarbons such as alkanes are the main constituents of the oily masses during spills and the degradation of the oily film is a fundamental step to prevent anoxia along the water column. The mechanisms of microbial aerobic oxidation can activate long-chain alkanes through the hydroxylation of the C-H bond via monooxygenase or dioxygenase enzymes: The alcohols, aldehydes and ketones that are derived from them are subsequently shunted towards the  $\beta$ -oxidation of fatty acids (Callaghan et al., 2006). P221-PWY, in which octane is oxidized up to octanoyl-acyl-CoA to enter fatty acid  $\beta$ -oxidation and which employs alkane monooxygenases, is highly enriched in polluted samples (Fig. 2C). Orthologs for the degradation of sulfur derivatives of methane such as dimethylsulfone monooxygenase (Fig. 2A, K17228) and metabolic pathways for the utilization of other short-chain hydrocarbons such as the P161-PWY for the degradation of acetylene were also detected in the contaminated samples (Fig. 2C, Table S2). Hydrocarbonoclastic enzymes and metabolic pathways involved in the degradation of a wide range of hydrocarbons were found to be enriched in the polluted samples: (i) naphthalene 1,2-dioxygenase ferredoxin component (K14578, EC: 1.18.1.7) for the degradation of PAHs, (ii) aromatic monocyclics derived from the oxidation of benzene PWY-5431 for the degradation of aromatic monocyclics derived from the oxidation of benzene, (iii) benzaldehyde dehydrogenase (K00141, EC: 1.2.1.28) involved in catechol degradation to  $\beta$ -ketoacid, (iv) Benzoate 1,2-dioxygenase (EC: 1.14.12.10), (v) 2'-hydroxybiphenyl-2-sulfinate desulfinate (EC: 3.13.1.3) and arylsulfatase (EC: 3.1.6.1) involved in the utilization of sulfure aromatics, (vi) 2-nitrobenzoate degradation metabolic pathway



**Fig. 2.** Hydrocarbonoclastic potential of lacustrine microbiomes. 16S gene sequence rRNA ASV profiles were used as inputs for PICRUSt2 to reconstruct metagenomic profiles in terms of KEGG orthologs and Enzyme Commission annotations. Metabolic pathway profiles were generated using MinPath and the MetaCyc pathway database. Gene counts for selected hydrocarbonoclastic genes were computed and compared for polluted and unpolluted samples. A. KEGG orthologs. B. Enzyme commission annotations (For description of KO and EC definitions, please refer to Supporting Information: Table S2). C. Abundances of selected metabolic pathways involved in the degradation of a range of hydrocarbon degradations was also graphed for polluted and unpolluted samples.

(PWY-5647) for the degradation of monocyclic and polycyclic nitrogen aromatics and dye decolorizing peroxidase (K15733, E1. 11.1.19) (Fig. 2A-C).

Additional hydrocarbonoclastic genes that were found to be enriched in polluted samples include alcohol dehydrogenase (EC: 1.1.1.1) and oxidoreductases (K00004). These enzymes show activity linked to aliphatic and aromatic primary and secondary alcohols (EC: 1.1.1.1 alcohol dehydrogenase propanol-preferring, EC: 1.1.1.90 Aryl-alcohol dehydrogenase). Similarly, cyclohexanone monooxygenase (EC: 1.14.13.22) is involved in the oxidation of cycloalcohol. All these reactions lead to the formation of aldehydes. Furthermore, metabolic pathways for fatty acid salvage or  $\beta$ -oxidation of fatty acids were also detected in polluted microbiomes (PWY-7094, Fig. 2C). PWY-7094 represents a central pathway for the aerobic degradation of aliphatic and aromatic hydrocarbons before the introduction of catabolites into the tricarboxylic acid cycle (Das and Chandran, 2011). Observed catalase (EC: 1.11.1.6), thanks to its huge oxygen production via the splitting of hydrogen peroxide molecules (one million molecules per enzyme), could assist in the costly oxidation processes of hydrocarbons (Fig. 2B) (Kaushal et al., 2018).

The predictive metagenomic analysis also highlighted the presence of multiple genes involved in the synthesis of biosurfactants such as peptidoglycan glycosyltransferase (EC 2.4.1.129, Fig. 2B, Table S2), N-

acetylglucosaminyl-diphospho-decaprenol L-rhamnosyltransferase (EC: 2.4.1.289), dTDP- 4-dehydrorhamnose reductase (EC:1.1.1.133), and phosphomannomutase (EC: 5.4.2.8). Phosphoglucomutase participates in rhamnolipid production, presumably catalyzing the first step in the deoxy-thymidine-diphospho-L-rhamnose (dTDP-L-rhamnose) pathway, (Olvera et al., 1999). Lipase (EC: 3.1.1.-), which hydrolyzes triglycerides into diglycerides and subsequently into monoglycerides and free fatty acids, was also enriched in polluted samples. The microbial extracellular lipase enzyme is highly soluble in water and acts at the surface of oil droplets to break down insoluble substrates that can be more easily absorbed (Marchut-Mikolajczyk et al., 2020). Overall, the predictive metagenomic analysis of Pertusillo communities demonstrated the hydrocarbonoclastic potential of the lacustrine microbiome in degrading complex substances such as crude oil. The metabolic potential for the degradation of aliphatic and aromatic sulfur, nitrogen and chlorinated hydrocarbons were evidenced through the detection of genes and pathways for emulsification and facilitation of degradation processes. Importantly, many enzymes involved in the degradation of sulfur hydrocarbons were detected in the polluted samples (K17228, K05977, EC: 1.14.14.5, EC: 3.1.6.1) and indeed, in Lake Pertusillo. (Fig.2A-C). Sulfur is the most abundant element in petroleum after carbon and hydrogen with a content that ranges from 0.03 to 7.89 mass% in crude oil (Solimani et al., 2007). The removal of sulfur from oil is one of the central

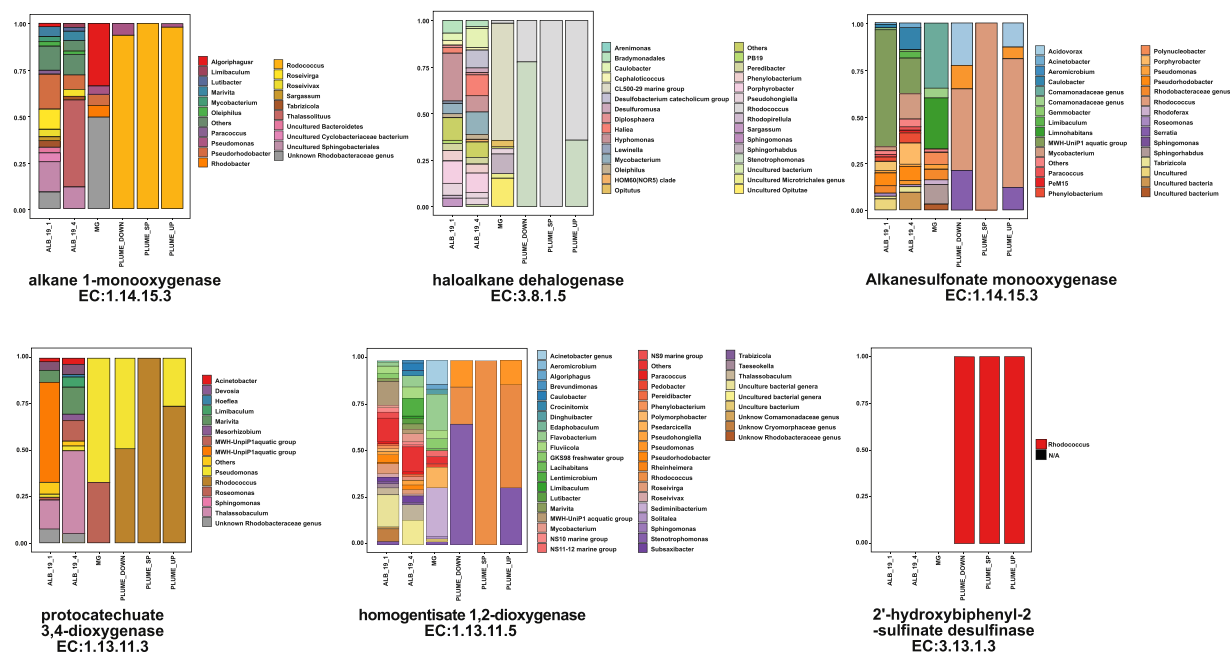


Fig. 3. Taxonomic contributions for various hydrocarbonoclastic genes across all samples determined at genus level with PICRUSt2.

requirements in most refineries and the price (and processing cost) of crude oil is often dictated by its sulfur content. Organic sulfur compounds such as thiols, sulfides, and thiophenic compounds represent the main source of sulfur found in crude oil. Sulfur contained in aromatic rings, such as thiophene and its benzo analogs are more resistant to sulfur removal by hydrodesulfurization and thermal conversion (Gray et al., 1995). The 2'-Hydroxybiphenyl-2-sulfinate desulfinase (K05977, DszB) enzyme found in the microbiomes of the polluted samples represent the terminal step of a catabolic pathway that converts dibenzothiophene (DBT), a common crude oil contaminant, into 2-hydroxybiphenyl (HBP) where a selective cleaving of carbon-sulfur bonds occurs.

### 3.4. Taxonomic contribute to oil degradation

To understand the primary contributors of hydrocarbon degradation potential in the Lake Pertusillo microbiome and in pristine samples, taxonomic clades contributing to genes for degradation of hydrocarbon and hydrocarbon derivatives were determined at genus level with PICRUSt2 (Table SI 2). Representative examples of taxonomic attribution to hydrocarbonoclastic genes for specific hydrocarbons are shown in Fig. 3. The general trend observed in pristine samples of Lake Alboraj and Lake Pertusillo (MG) is a broad taxonomic distribution for hydrocarbonoclastic enzymes such as monoxygenases. This diversity of genera is very restricted in contaminated samples where probably a selection due to the difficult environmental conditions has a substantial impact on biodiversity by converging the ability of these microorganisms to survive together with their hydrocarbonoclastic competence directed towards halogenated, sulfur and nitrogenous hydrocarbons (aliphatic and aromatics, Fig. 3). In this regard, the taxonomic contribution for desulfurisation of 2'-hydroxyphenyl-2-sulfinate (dszB; EC: 3.13.1.3) comes only from the genus *Rhodococcus*. This enzyme had very low counts in uncontaminated samples (Fig. 3, Table S2). DszB provides sulfate support for the microbial community by desulfurising recalcitrant and undesirable molecules found in crude oil such as dibenzothiophene (DBT) (Duarte et al., 2001). Rare enzymes that degrade resistant substances such as dibenzothiophene (DBT) could represent a selective driver that favours the success of *Rhodococcus* in these environments. The presence of desulfurization enzymes such as DszB and alkanesulfonate monoxygenase (SsuD, Fig. 3) not only indicate the

metabolic plasticity of *Rhodococcus*, but also makes them excellent candidates for the sustainable bio-desulfurization of crude oil (Peng et al., 2019; Táncsics et al., 2015; Viesser et al., 2020; Abbasian et al., 2016).

## 4. Discussion

There have been few remote sensing studies on oil spills in water bodies and reservoirs, whose waters are used for drinking purposes or other human and animal uses.

The field campaign, funded by PDVSA (Petróleos de Venezuela, S.A), aimed at developing an observation system able to detect and monitoring the oil spills in the lake Maracaibo. Only a limited number of studies have been conducted in ecosystems threatened by industrial accidents, such as Pertusillo and industrial oily wastewaters, like the Al Borma and Zenaiga desert reservoirs (Laneve et al., 2022). The online availability of data from the Copernicus platform means that observational studies of the Earth, even retrospective ones, can be carried out. Thus, the VNRI algorithm, used for the detection of oil pollution events in lake ecosystems, has shown that it can be used in real time and retrospectively for the monitoring of pollution events.

The genera *Pseudomonas*, *Serratia*, *Acidovorax*, *Stenotrophomonas* and *Rhodococcus* identified in the FOC areas of the oil polluted Lake Pertusillo were primary contributors to the hydrocarbonoclastic properties of the microbiome. These genera have been previously characterized during oil spills, which makes them hypothetically good biosensors of oil pollution (Mukherjee et al., 2017). Currently, only faecal pathogenic microorganisms are accepted and included in international regulations as indicators of contamination (Sagova-Mareckova et al., 2021). Increasing data are however available for biosensor microorganisms of different types of contamination, and this could offer further information on the quality of freshwater (for human use) that is threatened by contaminations (Sagova-Mareckova et al., 2021). In Lake Pertusillo, which is surrounded by a dense network of oil extraction plants, a hydrocarbonoclastic genus like *Rhodococcus* is abundant during an oil spill.

Taxonomic attributions identify the genus *Rhodococcus* as the only potential degrader of persistent sulfur substances involved in environmental pollution such as dibenzothiophene (DBT). Rare enzymes such as 2-hydroxybiphenyl-2-sulfinate desulfinase involved in the degradation

of DBT are assigned in this study to the genus *Rhodococcus* together with enzymatic patterns for the degradation of substances dangerous for human and environmental health such as aliphatic and aromatic hydrocarbons. The biodesulfurisation of dibenzothiophene and its derived compounds attracts the attention of the industrial world due to the lack of sustainable and economical techniques for the removal of organic sulfur. Industrially adopted methods are based on energy-intensive industrial processes of pyrolysis of petroleum substances.

The isolation of degrading microorganisms from aquifer ecosystems characterized by significant presence of polluting industrial plants could represent a valid means to combat local pollution and support industries for the *in situ* development of sustainable degradation processes for recalcitrant substances.

The information provided by satellite algorithmic maps combined with the metabolic reconstruction of resident communities could drive the isolation of species of interest for the bioremediation of water basins (Liu et al., 2022 May 12).

Similar to this study's approach to oil spills, the combination of satellite and metagenomic data in other environments can provide insight into ongoing ecosystem change. A study integrated satellite and metagenomic data for monitoring the effects of climate change in the Alpine and Polar Regions. The authors point out that the Arctic red snow phenomenon of Franz Josef Land is the dominant effect in the microbial population of the presence of *Chlamydomonas nivalis*, a red-colored unicellular photosynthetic green algae. White snow communities compared with red snow communities suggested that white snow were first colonized by fungal and virus and subsequently followed by *C. nivalis* present in red snow. Spectral reflectance satellite images showed that red snow covers up to 80% of the surface of the snow and ice fields in Franz Josef Land. *C. nivalis* supports a local food web that is expanding with increasing temperatures with widespread potential impacts on Alpine and polar environments worldwide (Hisakawa et al., 2015). In Brazil, metagenomics and satellites monitored marine impacts of iron tailings after the Fundão Dam rupture, which released 50 million m<sup>3</sup> of tailings into the Doce River. Seventeen days after the disaster the mud spread along the river and reached the coral reef of the Arolhos reef. Although, the coral reef microbiome do not highlighted substantial changes, these integrate methods could represent a support for these ecosystems, and be useful for preparing recovery plans for anthropogenic disturbances in the environment (Francini-Filho et al., 2019).

In this study, we demonstrate that the sequential application of multispectral algorithmic observations with omic sciences like metagenomic surveys can give indications on polluted areas and the communities that reside there. Detected microbial sensors of pollution may be used in the near future as new and effective indicators in environmental monitoring.

## 5. Conclusions

Sentinel2A (S-2A) and Sentinel-2B (S-2B) as part of the European Copernicus Program offer high levels of optical resolution and are equipped with a multi-spectral instrument (MSI) that divides the solar radiation reflected between 0.440 and 0.2200 μm into 13 multispectral channels. Mapping water bodies using satellite imagery plays an emerging role in water management. Research and development of algorithms to understand water surface changes in relation to the threat of dangerous pollutants is a satellite research challenge.

The sequential application and combination of multispectral algorithmic observations and metagenomic surveys can give indications on water polluted areas and the communities that reside there.

The pollution sensors identified in the waters of Pertusillo abound in typical catabolic patterns (e.g. alkane monooxygenase) suitable for the degradation of oily substances characteristic of oil spills such as aliphatic hydrocarbons and sulphurous hydrocarbon substances. Thus, genera as *Pseudomonas*, *Serratia*, *Acidovorax*, *Stenotrophomonas* and *Rhodococcus* (Mukherjee et al., 2017) identified in the FOC areas may

be used in the near future as new and effective indicators in environmental monitoring.

The microbial isolation from the microbiome of the areas identified by satellite can provide valid tools for the protection of the drinking water basins that are threatened by oil spills. The recurrence and the abundance of *Rhodococcus* sp. in Pertusillo waters and in different polluted environments (de Carvalho and da Fonseca, 2005) suggests a role of this bacteria during oil pollution. Metagenomic data and metabolic reconstruction suggest important role of *Rhodococcus* genus in Lake Pertusillo microbiome. Thus, due to the strong hydrocarbonclastic tendency of the genus *Rhodococcus* together with its rare tropism towards sulfur compounds such as DBT, the *Rhodococcus* sp. isolated in other studies (de Carvalho and da Fonseca, 2005) showed efficient catabolic machineries and could be valuable bioremediation microorganisms in oil spills (bioaugmentation *in situ*) and sustainable industrial processes such as petrol biodesulfurisation.

## Funding

This research was supported by the grant ISS: 4 ISS20-cd5d8b022b4e.

## CRedit authorship contribution statement

**Emilio D'Ugo:** Conceptualization, Methodology, Formal analysis, Writing – review & editing. **Arghya Mukherjee:** Software, Investigation, Formal analysis, Data curation. **Roberto Giuseppetti:** . **Matteo Tucci:** Software. **Paola Buccì:** Methodology. **Federico Aulenta:** Methodology, Formal analysis. **Giovanni Laneve:** Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Fabio Magurano:** Data curation, Conceptualization.

## Declaration of competing interest

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

## Data availability

Data will be made available on request.

## Acknowledgments

We would like to thank Milena Bruno, Valentina Messineo, Giorgio Santoriello for their assistance and Cova Contro onlus, via Longarone 24, Policoro (MT), associazionecovacontro@gmail.com. Movimento Libieriamo la Basilicata, Piazza Crispi 1, Potenza, liberiamolabasilicata@pec.it.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.envadv.2024.100498.

## References

- Sousa, João C.G., Ribeiro, Ana R., Barbosa, Marta O., Fernando R. Pereira, M., Silva, Adrián M.T., 2018. A review on environmental monitoring of water organic pollutants identified by EU guidelines. *J. Hazard. Mater.* 344, 146–162. <https://doi.org/10.1016/j.jhazmat.2017.09.058>. ISSN 0304-3894.
- Bramich, J, Bolch, CJS, Fischer, A., 2021. Improved red-edge chlorophyll-a detection for Sentinel 2. *Ecol. Indic.* 120, 106876.
- Drusch, M, Del Bello, U, Carlier, S, Colin, O, Fernandez, V, Gascon, F, et al., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens. Environ.* 120, 25–36.



- Rodríguez-Benito, CV, Navarro, G, Caballero, I., 2020. Using Copernicus Sentinel-2 and Sentinel-3 data to monitor harmful algal blooms in Southern Chile during the COVID-19 lockdown. *Mar. Pollut. Bull.* 161, 111722.
- Kolokoussis, P, Karathanassi, V., 2018. Oil spill detection and mapping using sentinel 2 imagery. *J. Mar. Sci. Eng.* 6, 4.
- Laneve, G, Bruno, M, Mukherjee, A, Messineo, V, Giuseppetti, R, De Pace, R, et al., 2022. Remote sensing detection of algal blooms in a lake impacted by petroleum hydrocarbons. *Remote Sens.* 14.
- Andreato, FD, Jiménez, DJ, Chaves, D, Dias, ACF, Luvizotto, DM, Dini-Andreato, F, et al., 2012. The microbiome of Brazilian mangrove sediments as revealed by metagenomics. *PLoS. One* 7, e38600.
- Mukherjee, A, Chettri, B, Langpoklakpam, JS, Basak, P, Prasad, A, Mukherjee, AK, et al., 2017. Bioinformatic approaches including predictive metagenomic profiling reveal characteristics of bacterial response to petroleum hydrocarbon contamination in diverse environments. *Sci. Rep.* 7, 1108.
- Yang, S, Wen, X, Jin, H, Wu, Q., 2012. Pyrosequencing Investigation into the bacterial community in permafrost soils along the China-Russia crude oil pipeline (CRCOP). *PLoS. One* 7, e52730.
- D'Ugo, E, Bruno, M, Mukherjee, A, Chattopadhyay, D, Giuseppetti, R, De Pace, R, et al., 2021. Characterization of microbial response to petroleum hydrocarbon contamination in a lacustrine ecosystem. *Environ. Sci. Pollut. Res. Int.* 28, 26187–26196.
- Mason, OU, Hazen, TC, Borglin, S, Chain, PSG, Dubinsky, EA, Fortney, JL, et al., 2012. Metagenome, metatranscriptome and single-cell sequencing reveal microbial response to Deepwater Horizon oil spill. *ISME J.* 6, 1715–1727.
- Sagova-Mareckova, M, Boenigk, J, Bouchez, A, Cermakova, K, Chonova, T, Cordier, T, Eisendle, U, Elserk, T, Fazi, S, Fleituch, T, Frühe, L, Gajdosova, M, Graupner, N, Haegerbaeumer, A, Kelly, AM, Kopecky, J, Leese, F, Nöges, P, Orlic, S, Panksep, K, Pawlowski, J, Petrussek, A, Piggott, JJ, Rusch, JC, Salis, R, Schenk, J, Simek, K, Stovicek, A, Strand, DA, Vasquez, MI, Vrålstad, T, Zlatkovic, S, Zupancic, M, Stoeck, T., 2021. Expanding ecological assessment by integrating microorganisms into routine freshwater biomonitoring. *Water. Res.* 191, 116767 <https://doi.org/10.1016/j.watres.2020.116767>. Mar 1Epub 2020 Dec 23. PMID: 33418487.
- Smith, MB, Rocha, AM, Smillie, CS, Olesen, SW, Paradis, C, Wu, L, et al., 2015. Natural bacterial communities serve as quantitative geochemical biosensors. *mBio* 6, e00326. -15.
- dos Santos, HF, Cury, JC, do Carmo, FL, dos Santos, AL, Tiedje, J, van Elsas, JD, et al., 2011. Mangrove bacterial diversity and the impact of oil contamination revealed by pyrosequencing: bacterial proxies for oil pollution. *PLoS One* 6, e16943.
- D'Ugo, E, Marcheggiani, S, Fioramonti, I, Giuseppetti, R, Spurio, R, Helmi, K, et al., 2016. Detection of human enteric viruses in freshwater from European countries. *Food Environ. Virol.* 8, 206–214.
- Douglas, GM, Maffei, VJ, Zaneveld, JR, Yurgel, SN, Brown, JR, Taylor, CM, et al., 2020. PICRUSt2 for prediction of metagenome functions. *Nat. Biotechnol.* 38, 685–688.
- Caspi, R., et al., 2020. The MetaCyc database of metabolic pathways and enzymes - a 2019 update. *Nucleic. Acids. Res.* 48, D445–D453. <https://doi.org/10.1093/nar/gkz862>.
- Wickham, H, Chang, W, Wickham, MH., 2016. Package 'ggplot2'. Create elegant data visualisations using the grammar of graphics. Version 2, 1–189.
- Eze, MO, Hose, GC, George, SC, Daniel, R., 2021. Diversity and Metagenome Analysis of a Hydrocarbon-Degrading Bacterial Consortium from Asphalt Lakes Located in Wietze, 11. *AMB Express, Germany*, p. 89. <https://doi.org/10.1186/s13568-021-01250-4>.
- Callaghan, AV, Gieg, LM, Kropp, KG, Suflija, JM, Young, LY., 2006. Comparison of mechanisms of alkane metabolism under sulfate-reducing conditions among two bacterial isolates and a bacterial consortium. *Appl. Environ. Microbiol.* 72, 4274–4282.
- Das, N, Chandran, P., 2011. Microbial degradation of petroleum hydrocarbon contaminants: an overview. *Biotechnol. Res. Int.* 2011, 941810.
- Kaushal, J, Mehandia, S, Singh, G, Raina, A, Arya, SK., 2018. Catalase enzyme: Application in bioremediation and food industry. *Biocatal. Agric. Biotechnol.* 16, 192–199.
- Marchut-Mikolajczyk, O, Drożdżyński, P, Struszczyk-Świta, K., 2020. Biodegradation of slop oil by endophytic *Bacillus cereus* EN18 coupled with lipase from *Rhizomucor miehei* (Palatase®). *Chemosphere* 250, 126203.
- Soleimani, M, Bassi, A, Margaritis, A., 2007. Biodesulfurization of refractory organic sulfur compounds in fossil fuels. *Biotechnol. Adv.* 25, 570–596.
- Gray, MR, Ayasse, AR, Chan, EW, Veljkovic, M., 1995. Kinetics of hydrodesulfurization of thiophenic and sulfide sulfur in athabasca bitumen. *Energy Fuels* 9, 500–506.
- Duarte, GF, Rosado, AS, Seldin, L, de Araujo, W, van Elsas, JD., 2001. Analysis of bacterial community structure in sulfurous-oil-containing soils and detection of species carrying dibenzothiophene desulfurization (dsz) genes. *Appl. Environ. Microbiol.* 67 (3), 1052–1062. <https://doi.org/10.1128/AEM.67.3.1052-1062.2001>.
- Peng, C, Huang, D, Shi, Y, Zhang, B, Sun, L, Li, M, Deng, X, Wang, W., 2019. Comparative transcriptomic analysis revealed the key pathways responsible for organic sulfur removal by thermophilic bacterium *Geobacillus thermoglucosidasius* W-2. *Sci. Total Environ.* 676, 639–650.
- Táncsics, A, Benedek, T, Szoboszlai, S, Veres, PG, Farkas, M, Máthé, I, et al., 2015. The detection and phylogenetic analysis of the alkane 1-monoxygenase gene of members of the genus *Rhodococcus*. *Syst. Appl. Microbiol.* 38, 1–7.
- Viesser, JA, Sugai-Guerios, MH, Malucelli, LC, Pincerati, MR, Karp, SG, Maranhão, LT., 2020. Petroleum-tolerant rhizospheric bacteria: isolation, characterization and bioremediation potential. *Sci. Rep.* 10, 2060.
- Abbasian, F, Palanisami, T, Megharaj, M, Naidu, R, Lockington, R, Ramadass, K., 2016. Microbial diversity and hydrocarbon degrading gene capacity of a crude oil field soil as determined by metagenomics analysis. *Biotechnol. Prog.* 32, 638–648.
- Liu S, Moon CD, Zheng N, Huws S, Zhao S, Wang J. Opportunities and challenges of using metagenomic data to bring uncultured microbes into cultivation. *Microbiome.* 2022; 10(1):76. [10.1186/s40168-022-01272-5](https://doi.org/10.1186/s40168-022-01272-5). PMID: 35546409; PMCID: PMC9097414.
- Hisakawa, N., Quistad, S.D., Hester, E.R., 2015. Metagenomic and satellite analyses of red snow in the Russian Arctic. *PeerJ.* 3, e1491.
- Francini-Filho, R.B., Cordeiro, M.C., Omachi, C.Y., Rocha, A.M., Bahiense, L., Garcia, G. D., Tschoeke, D., de Almeida, M.G., Rangel, T.P., De Oliveira, B.C.V., et al., 2019. Remote sensing, isotopic composition and metagenomics analyses revealed Doce River ore plume reached the southern Abrolhos Bank Reefs. *Sci. Total Environ.* 697, 134038.
- de Carvalho, C.C.C.R., da Fonseca, M.M.R., 2005. The remarkable *Rhodococcus erythropolis*. *Appl. Microbiol. Biotechnol.* 67, 715–726. <https://doi.org/10.1007/s00253-005-1932-3>.